



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

**DESIGN AND ANALYSIS OF AN ELECTRON
GUN/BOOSTER AND FREE ELECTRON LASER
OPTICAL THEORY**

by

Sean P. Niles

September 2010

Dissertation Supervisor:

William B. Colson

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, Va 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY (<i>Leave blank</i>)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED	
	September 2010	Dissertation	
4. TITLE AND SUBTITLE Design and Analysis of an Electron Gun/Booster and Free Electron Laser Optical Theory		5. FUNDING NUMBERS	
6. AUTHORS Niles, Sean P.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol Number:_____			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		12b. DISTRIBUTION CODE	
13. ABSTRACT(<i>maximum 200 words</i>) As interest in high power free electron lasers (FELs) has increased, the FEL and accelerator communities have been faced with the need to develop high bunch charge, high repetition rate, low emittance electron sources for use as the driving accelerators for FELs. A novel superconducting, radio-frequency (SRF) gun/booster has been designed by and built for the Naval Postgraduate School (NPS) FEL Beam Physics Lab in collaboration with Niowave, Inc., for studying this electron source regime. The NPS SRF gun/booster operates at 500 MHz and is based upon a quarter-wave structure. It incorporates many features that make it desirable for studying the cathodes and transport regimes necessary to explore high bunch charge beams, including adjustable field focusing, short transport out of the gun, and the ability to change cathode types and materials. After attaining "first beam" in June 2010, the NPS gun has been established as the first SRF electron gun in the United States. Initial results show excellent agreement with simulation with bunch charges of 110 pC and transverse emittance estimates of ~4 mm-mrad. Additionally, a modal analysis tool for the NPS FEL simulation software is developed based upon the Hermite-Gaussian basis set. Using a minimization of mode coefficients approach, we decompose output optical fields for amplifier FEL designs and experiments for FEL optimization and comparison of laser output fields.			
14. SUBJECT TERMS Free Electron Laser, Electron Gun, Quarter Wave, Electron Accelerator, Hermite-Gaussian, Optical Decomposition		15. NUMBER OF PAGES 263	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

**DESIGN AND ANALYSIS OF AN ELECTRON GUN/BOOSTER AND
FREE ELECTRON LASER OPTICAL ANALYSIS**

Sean P. Niles

Lieutenant Commander, United States Navy
B.N.E., Georgia Institute of Technology, 1997
M.S., Naval Postgraduate School, 2005

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN PHYSICS

from the

NAVAL POSTGRADUATE SCHOOL
September 2010

Author:

Sean P. Niles

Approved by:

William B. Colson
Professor of Physics
Dissertation Supervisor

John Lewellen
Professor of Physics

Todd Smith
Professor of Physics
Stanford University

Robert Armstead
Professor of Physics

James H. Newman
Professor of Space Systems

Joseph Blau
Professor of Physics

Approved by:

Andres Larraza, Chairman, Department of Physics

Approved by:

Doug Moses, Vice Provost for Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

As interest in high power free electron lasers (FELs) has increased, the FEL and accelerator communities have been faced with the need to develop high bunch charge, high repetition rate, low emittance electron sources for use as the driving accelerators for FELs. A novel superconducting, radio-frequency (SRF) gun/booster has been designed by and built for the Naval Postgraduate School (NPS) FEL Beam Physics Lab in collaboration with Niowave, Inc., for studying this electron source regime. The NPS SRF gun/booster operates at 500 MHz and is based upon a quarter-wave structure. It incorporates many features that make it desirable for studying the cathodes and transport regimes necessary to explore high bunch charge beams, including adjustable field focusing, short transport out of the gun, and the ability to change cathode types and materials. After attaining “first beam” in June 2010, the NPS gun has been established as the first SRF electron gun in the United States. Initial results show excellent agreement with simulation with bunch charges of 110 pC and transverse emittance estimates of \sim 4 mm-mrad. Additionally, a modal analysis tool for the NPS FEL simulation software is developed based upon the Hermite-Gaussian basis set. Using a minimization of mode coefficients approach, we decompose output optical fields for amplifier FEL designs and experiments for FEL optimization and comparison of laser output fields.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	ELECTRON SOURCES AND DRIVE LASER	5
	A. CATHODES OF INTEREST	5
	1. Thermionic Cathodes	5
	2. Metal and Semiconductor Cathodes	8
	3. Field Emission Cathodes	10
	B. DESIGN OF A UV DRIVE LASER SYSTEM	11
	1. Minimalist Approach	12
	2. Optimized Design	12
III.	SRF GUN/BOOSTER	15
	A. RF CAVITY TYPES	20
	1. Right Cylinder Pillbox Cavity	20
	2. Tesla Cavity	24
	3. Quarter-Wave Cavity	26
	B. FABRICATION	30
	1. Cavity	32
	2. Coupler	38
	3. Cathode	43
	4. Superconducting Solenoid	52
	5. Assembly and Cavity Cleaning	53
IV.	MODELING AND EXPERIMENT	83
	A. MODELING	83
	1. Simulation Codes Used	84
	2. Beam Simulation Methodology	91
	3. Space Charge	93
	4. DC Gun Analysis	101

5.	Booster Configuration	114
6.	Gun Configuration	122
7.	Injector Proposal	134
B.	SRF GUN/BOOSTER EXPERIMENTS	150
1.	Resonant Modes	150
2.	Bead Pull	155
3.	SC Solenoid	159
4.	Cold Cavity Processing	172
5.	Full Beam Test	175
V.	FEL BACKGROUND	187
A.	HISTORY AND COMPONENTS	187
1.	History	187
2.	FEL Components	189
B.	BASIC FEL PHYSICS	192
1.	Lorentz Equations	192
2.	Electron Motion in a Periodic Magnetic Field	192
3.	Resonance Condition	194
4.	Current Effects	195
VI.	HERMITE-GAUSSIAN DECOMPOSITION	199
A.	HG SOLUTIONS TO THE WAVE EQUATION	200
1.	The Electromagnetic Wave Equation	200
2.	HG Solutions	203
3.	Higher Order HG Modes	206
B.	FIELD DECOMPOSITION	213
C.	CHARACTERIZING FEL FIELDS	217
VII.	CONCLUSION	231
LIST OF REFERENCES		235
INITIAL DISTRIBUTION LIST		241

LIST OF FIGURES

1.	A simple diode	6
2.	A simple triode	7
3.	Scanning electron micrograph of a diamond field emitter	11
4.	Stanford test stand laser	13
5.	Cathode test stand laser	14
6.	Stanford superconducting accelerator DC gun	16
7.	Rossendorf superconducting electron gun	19
8.	Pillbox cavity	20
9.	Single-point multipacting trajectories in a muffin-tin cavity	25
10.	HEPL cavities	26
11.	Tesla cavities	27
12.	Tesla cavity multipacting trajectories	28
13.	A quarter-wave cavity	29
14.	NPS 500 MHz superconducting gun/booster	31
15.	Quarter-wave pillbox cavity	33
16.	Quarter-wave pillbox cavity dimensional effects	34
17.	NPS axial cavity fields as a function of cathode position	35
18.	NPS radial cavity fields as a function of cathode position	36
19.	NPS cavity base parts	37
20.	NPS cavity RF coupler drawing	39
21.	Cavity coupling constants	40
22.	Transverse coaxial coupler	42
23.	Simplified cathode assembly model	45
24.	Rossendorf and BNL RF chokes	48
25.	NPS cavity cathode assembly	49
26.	NPS cavity cathode tip detail	51

27.	Superconducting solenoid bus bars	52
28.	Electron beam welding of the NPS cavity	54
29.	Buffered chemical polish station	55
30.	High-pressure rinse station	57
31.	NPS SRF cavity leak check post-chemical processing	58
32.	Cavity temperature monitoring	59
33.	Cavity foundation	60
34.	Insulating blanket material	61
35.	Liquid nitrogen shield	62
36.	MuMetal shield	63
37.	Vacuum vessel assembly	64
38.	Conning tower assembly	64
39.	Warm cavity supports	65
40.	Coupler and pump station	66
41.	RF coupler assembly	67
42.	Coupler insertion	68
43.	RF testing setup	69
44.	Assembly complete	70
45.	Portable clean room and cathode assembly	71
46.	Teflon rails under cathode assembly structure	72
47.	Preparations prior to opening the cathode and cryomodule systems	73
48.	Cathode assembly spider damage	74
49.	Prepared to open the cavity	75
50.	Insertion of the cathode assembly	76
51.	A PVC pusher is used to position the spider during insertion.	77
52.	Final mating of the cathode assembly to the cavity	78
53.	Optical alignment of the cathode assembly	79
54.	Major registration marks in the interior of the diagnostic beamline	80

55.	Laser spots on the cathode	81
56.	Current system for envelope equation	94
57.	Uniform-Uniform bunch distribution	96
58.	Gaussian-Gaussian bunch distribution	97
59.	Non-linear, asymmetric bunch distribution	98
60.	Space charge effects comparison	99
61.	DC gun voltage divider	102
62.	First beam at the NPS Beam Physics Lab	103
63.	DC gun diagnostic line	104
64.	DC gun diagnostic line Faraday cups	105
65.	Stanford DC gun equipotentials	106
66.	Energy plots for various electrode ratios	107
67.	Envelope of beam for various electrode ratios	108
68.	QE vs. cathode heater voltage	110
69.	QE as a function of decay time	111
70.	Eimac cathode diagram	113
71.	Acceptance of NPS cavity as booster	115
72.	Emittance and radius dependence on phase for the NPS cavity as booster	117
73.	Transport through NPS cavity as booster	118
74.	Transport through NPS cavity as booster with solenoid at input drift .	119
75.	Emittance comparison for booster configuration with and without input solenoid	121
76.	500 MHz gun diagnostic beam line	123
77.	Acceptance of NPS cavity in gun configuration	125
78.	Radius and emittance of NPS cavity as gun phase is changed	126
79.	Acceptance of various beam shapes by the NPS gun	128
80.	Radius and emittance of various beam shapes in the NPS gun as gun phase is changed	129

81.	Beam envelopes of various beam shapes from the NPS gun	132
82.	Emittance evolution of various beam shapes from the NPS gun	133
83.	Double bell quarter wave booster cavity	134
84.	Long double QW booster cavity electric fields	137
85.	Short double QW booster cavity electric fields	138
86.	3D long booster cavity acceptance for emittance optimization	141
87.	2D long booster cavity acceptance for spot emittance optimization . . .	142
88.	3D short booster cavity acceptance plot for spot size optimization . .	143
89.	2D short booster cavity acceptance for spot size optimization	144
90.	Beam envelope of long and short booster cavities	146
91.	Transverse emittance evolution of long and short booster cavities . .	147
92.	Final longitudinal emittance long and short booster cavities	149
93.	COMSOL field map of the NPS booster cavity	152
94.	NPS gun/booster cavity configuration for eigenfrequency testing . . .	153
95.	Network analyzer plots of NPS gun/booster cavity eigenfrequencies .	154
96.	Bead pull setup for NPS cavity	156
97.	Bead pull phase shift data from NPS booster cavity measurements . .	157
98.	Bead pull booster cavity electric field profile	158
99.	Superfish solenoid magnetic field map	160
100.	COMSOL solenoid magnetic field map	161
101.	Experimental set up for measuring on axis magnetic field of a solenoid	163
102.	Magnetic field map data	164
103.	Inspecting the superconducting solenoid after removal from the liquid helium bath.	165
104.	Inductance data of superconducting solenoid	166
105.	Simulation of beam rotation through the superconducting solenoid . .	169
106.	Q_0 vs. E plot for the NPS booster cavity	174
107.	Cathode position impact on f_0	176

108.	Instrumentation stations in the NPS gun diagnostic beam line	178
109.	Beam cross sections	179
110.	Window-frame steering corrector	180
111.	Beam energy estimation using window-frame correctors	181
112.	Faraday cup bunch charge signals	183
113.	NPS gun relative phase acceptance	185
114.	System layout diagram of the JLAB FEL	189
115.	Linear undulator diagram showing electron trajectory	191
116.	Hermite-Gaussian intensity patterns	212
117.	Area increase with mode number	219
118.	Mode composition based on scaling factor changes	220
119.	Coefficient area topology of a Gaussian mode	221
120.	Nearly Gaussian coefficient area topology maps	222
121.	Graphical Cross-correlation Demonstration	223
122.	2D cross-correlation surface	225
123.	Cross-correlation maps	227
124.	BNL SDL optical field amplitude	230
125.	BNL SDL Gain	230

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

1.	Cavity AC power requirements	17
2.	Rossendorf 650 MHz SRF gun parameters	18
3.	Ohmic heating of the NPS 500 MHz gun cathode assembly	50
4.	Long pulse vs. Short pulse space charge effects	100
5.	Longitudinal emittances for the NPS gun and proposed injector designs.	147
6.	NPS gun/booster cavity eigenfrequencies and comparison with simulation	154
7.	NPS gun bunch charge and QE measurements	184
8.	Hermite polynomials through 7th order	211
9.	FEL parameters used for HG decomposition.	228
10.	Modal analysis results of four amplifier FELs.	229

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Use of the free electron laser (FEL) has blossomed as an analysis tool for science, as well as developing research opportunities for possible high power applications. Current FELs are generally associated with large research institutions and are attached to electron accelerators and storage rings. As interest in high power has increased, the FEL and accelerator communities have been faced with the need to develop high bunch charge, high repetition rate, low emittance electron sources for use as the driving accelerators for the FELs. A novel superconducting, radio-frequency (SRF) gun/booster has been designed by and built for the Naval Postgraduate School (NPS) FEL Beam Physics Lab in collaboration with Niowave, Inc., for studying this electron source regime. This research supports the development of the NPS FEL and the U.S. Navy's FEL Innovative Naval Prototype.

The NPS SRF gun/booster operates at 500 MHz and is based upon a quarter-wave structure rather than the more prevalent elliptical cell design. It incorporates many features that make it desirable for studying the cathodes and transport regimes necessary to explore high bunch charge beams, including adjustable field focusing, short transport out of the gun, and the ability to change cathode types and materials. The cathode system allows for the use of conventional metal and semiconductor cathodes and the cavity generates fields of sufficient strength for testing field emission cathodes as well. With the cathode stalk removed, the cavity can be used as a booster cavity in an injector system with thermionic cathodes. Numerous simulations during the design phase are compared with experimental results and show excellent agreement. Experimentation and simulation by the author as part of the collaboration team is ongoing and results for bunch charge, beam energy, normalized emittance, and proposed improvements to date are presented and discussed.

The development and construction of the prototype cavity is also an excellent case study for rapid prototyping in the SRF accelerator field as the gun/booster cavity

proceeded from initial design to first cold-test in less 12 months, shattering previous SRF gun development cycles by years. It is the first operational SRF electron gun in the United States.

To develop a common comparison scheme for FELs and conventional lasers, a modal analysis tool for the NPS FEL simulation software is developed based upon the Hermite-Gaussian basis set. In amplifier FELs, there are no external factors determining the appropriate transverse size and phase scaling factors to specify the basis set for modal decomposition. Using a minimization of mode coefficients approach, the author demonstrates the ability to decompose output optical fields for various amplifier FEL designs and experiments as tool for FEL optimization and comparison of laser output optical fields.

ACKNOWLEDGMENTS

No work of this type is ever accomplished completely on one's own and this dissertation is no different. I received support from many over the course of my doctoral work, and I appreciate the time that was generously given in order to answer my questions, decipher my drawings, and sometimes just listen to my complaints.

At the top of my list is my family. Through the long nights sitting at the dining room table while typing or away from home at a conference or experimental run, my wife kept the household running. There are not words that can express how proud I am of my wife, but I know I married the right woman...how she puts up with me I will never know, but I am happy she does. To Lindsey and Madeline, I know you do not understand what "Daddy going to school" really means, but I hope that I can kindle the interest in science this process appears to have sparked. I look forward to having some more free time so we can get to all those really neat experiments you want to do.

The members of my committee have been outstanding. Professor Colson welcomed me back to the FEL group with open arms and encouraged me through some of the more difficult times during the process. The addition of Professor Lewellen to the research group opened the exciting door of experimental work, and I appreciate his patience while I learned how not to break things in the lab. Luckily everything was replaceable. Professors Armstead, Blau, Newman, and Smith were also extremely helpful, especially in the presentation and explanation of my work. To all, the ongoing discussions, comments and questions (and various colors of ink) helped me craft a much more concise and useful dissertation. Your patience and guidance are very much appreciated.

Professor Borden provided a much needed sounding board and his sage advice and encouragement were much appreciated. I wish I could take more of his classes (but I am sure he would prefer I did not). Even if he will not accept the mantle of

“mentor,” I think his teaching style and interaction with students is how I would like to model my future teaching career.

Without my collaborators at Niowave and Boeing, none of the experimental work on the SRF cavity would have been possible. I learned a lot about the hardware necessary for drive lasers and accelerators, and I appreciated everyone’s willingness to slow down some of the conversations so I could keep up. Niowave’s material support in construction and operation of the gun were instrumental to getting the gun into test as quickly as we did. Without the drive laser and experts provided by Boeing, we would have had a cavity, but no electrons to accelerate. The laser operated flawlessly and allowed us to focus on the gun behavior and electron transport. I look forward to working with all of you in the future, but will definitely try to schedule my visits for the warmer months of the year.

The professors and students of the FEL research group have had to put up with me for as long as 5 years between my two tours. I have many memories forged by our lively conversations and trips. Someday, I will have to work on the HO-scale railgun, and I will never forget watching 5 PhDs arguing with a barmaid over whether their pints really measured a fluid pint at the snug in Victoria. You have succeeded in turning someone who believed he was an engineer into someone who realizes that physicists get to work on all the really interesting questions and is glad to have changed his title.

To the other PhD students in the physics department, your commiseration has helped me immensely. It feels good to not be the only one that has struggled. Thanks for letting me vent when I needed to, and I look forward to hearing about your future successes!

I. INTRODUCTION

Free electron lasers (FEL) represent an active interdisciplinary research area incorporating challenges in materials, optical transport, electron sources, accelerator design, and many other areas of interest to the scientist and engineer. As the technology associated with the FEL has matured, its usefulness as a laser that can be designed for many applications has increased demand for these systems worldwide in both academia and industry.

While current user facilities are quite large and typically associated with large research universities or government user facilities, technology improvements are making many of the required subsystems smaller. There are applications in homeland defense, military settings, and other science fundamental and applied research programs that have an interest in smaller FEL designs. Particularly in military applications, there is also a desire to increase the output laser power of the FEL to increase its viability as a self-defense system. For many of these thrusts, necessary research in electron beam sources and linear accelerators is necessary.

In this work, we report on the building and testing of a new superconducting radio frequency (SRF) electron gun/booster. The gun/booster is the first quarter wave structure designed to accelerate electrons and is unique in that it can be used as an energy booster, or, with the addition of a cathode stalk, as an electron gun. The gun/booster is designed to operate at 500 MHz, below the frequency of many current RF guns in use. Due to the quarter wave structure, it is still extremely small compared to its RF wavelength—the entire cryomodule is less than \sim 1 m long.

The full cavity development to “first beam” cycle was accomplished in less than 24 months. The cavity’s design and manufacturing progress has been presented in numerous poster and oral presentations as well in conference papers by the author and other collaborators. This approach is considerably different than the standard cavity design approach in which a cavity design is proposed and redesigned multiple

times before the cavity is ever fashioned, sometimes many years later. Our approach is based on rapid prototyping—understanding that the initial design will have flaws, but in the belief that experience will provide invaluable insight, enabling later cavity generations to be far superior than if only paper studies were conducted. Cavity performance during the initial testing has been below design criteria, but experimental results look promising for this type of cavity design.

Since the end measure of effectiveness for an FEL is the output laser light, we develop a theoretical framework for modal analysis of the output light from an amplifier FEL. Based upon a Hermite-Gaussian mode decomposition of the optical field, an optimization of the free scaling parameters is performed to minimize the number of modes necessary to represent the optical field. By comparing the modal content of an FEL and a conventional laser, it is possible to perform a near “apples-to-apples” comparison of the expected propagation characteristics.

This FEL analysis tool developed and coded by the author has been incorporated into the Naval Postgraduate School (NPS) FEL4D simulation code and is used to analyze actual and proposed amplifier FEL designs as published by the author in collaboration with the the NPS FEL theory group [1]. Additionally, with planned improvements to the FEL code allowing the importation of electron distributions from common particle accelerator codes, it becomes possible to couple the performance of new gun and accelerator designs with the FEL simulation and analyze directly the impact of these changes.

In this dissertation, we will discuss briefly the cathode technologies that are compatible with current experiments using legacy equipment from the former Stanford FEL and for future use in the NPS SRF gun. The NPS gun/booster development will be traced and important design considerations discussed and analyzed. The details of the cavity construction process are documented for reference by future researchers. Beam simulation studies performed by the author for the NPS gun/booster in direct preparation for experimentation are presented, as well as a possible injector design

based upon the NPS SRF cavity configuration. Cold cavity and first beam tests developed by the author and collaboration team are conducted and presented showing that the gun design is successful and shows promise as a possible beam source for a linear accelerator. Finally, after a short FEL tutorial, we present the Hermite-Gaussian analysis tool and demonstrate its capabilities through FEL analysis.

THIS PAGE INTENTIONALLY LEFT BLANK

II. ELECTRON SOURCES AND DRIVE LASER

Generating an electron beam is essential in the operation of a free electron laser (FEL). The specific electron source can vary from machine to machine, as can the process for liberating those electrons from their source to be accelerated. In this chapter, we examine electron sources of interest to the research considered in this work as well in future work for the research group. We also describe the drive laser developed for cathode studies discussed later in this work.

A. CATHODES OF INTEREST

There are many ways to liberate electrons from a material to put into a beam. The three primary methods used in conjunction with linear accelerators are photoemission, thermionic emission and field emission. Some common desirable characteristics for a good electron source would be a low thermal emittance at launch, high current density, long lifetime, being robust, and providing a reliable method for beam shaping. In order of development, we will discuss the attributes of each type cathode and their applicability to our research.

1. Thermionic Cathodes

Thermionic cathodes have been in use since the first cathode ray tubes. In its simplest form, the thermionic cathode is a hot wire or plate in an electric field. The thermal energy imparted to the electrons provides them enough energy to escape the metal's surface. Once free of the metal, the electric field accelerates the electrons toward an anode. If the anode is a mesh grid or annular in shape, the electrons can flow through, creating a beam. A configuration with only a cathode and anode is referred to as a diode [2]. A simple diode is shown in Figure 1.

As the name implies, the thermionic cathode relies on thermal energy. If one desires to increase the current flow from a thermionic cathode, all other variables held constant, one must increase the cathode temperature. Obviously, this cannot be done

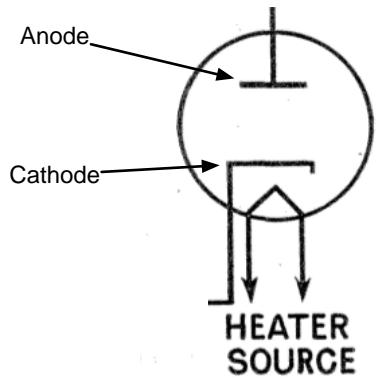


Figure 1. A simple diode drawing. Electrons are liberated from a cathode by heating and accelerated toward an anode maintained at a higher potential. After [3].

indefinitely, as the cathode will eventually melt. Unfortunately, changing the cathode temperature is not a fast process, so attempting to generate a high frequency pulsed beam requires altering other variables in the system. One can vary the cathode-anode voltage, thereby changing the field seen by the electrons freed by their thermal energy. When the field is applied, beam is formed; when the field is off, electrons are not pulled from the cathode and no beam is formed. Unfortunately for linear accelerators, the required cathode-anode voltage is typically in the hundreds of kV and switching high voltages at high speeds is not something to be undertaken lightly. There exists another, more tractable, method for generating a pulsed beam from a thermionic cathode.

If we add another electrode, again meshed so the beam can pass through it, near the cathode surface, we can apply a much lower voltage and overcome the higher fields provided by the main anode. A diagram of this arrangement, called a triode, is shown in Figure 2. This type of thermionic cathode is what was installed in the Stanford free electron laser injector that is currently being used by the Naval Postgraduate School (NPS) for cathode studies. Since the voltage applied to the grid

is small (hundreds of volts or less), it can be switched quickly to provide sharp on-off edges for the beam as well as allowing for a high repetition rate.

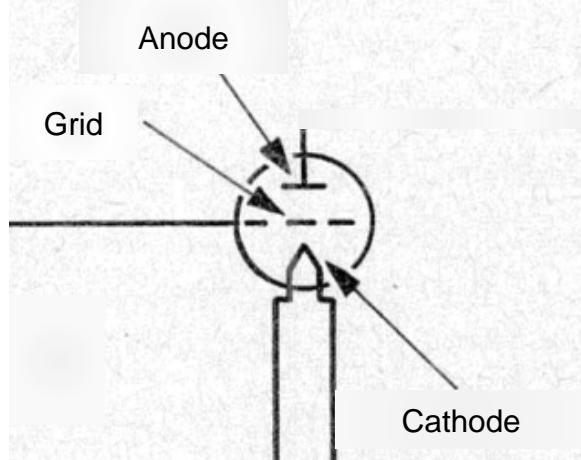


Figure 2. A simple triode. Electrons are liberated from the cathode and accelerated toward a higher potential anode. Between the cathode and anode is a gate electrode that can be biased to enhance or stop electron flow across the cathode-anode gap. After [3].

As these cathodes have developed since the late 1800s, many more improvements have been made. As it was discovered that impurities can alter the electron affinity of the cathode surface, manufacturers began developing “dispenser” cathodes to purposely put beneficial impurities onto the surface, greatly increasing their sourcing capabilities. Electrode geometries have also improved as it was realized that the electric field lines can be used to shape the beam as it leaves the cathode, for example the Pierce geometry [4]. Additionally, photo-switching now allows thermionic cathodes to be switched at extremely high rates. In photo-switching, the cathode is used as a thermally assisted photoemitter, only emitting when a laser shines on the cathode surface.

In general, thermionic cathodes are robust. For example, the cathode installed in the Stanford gun currently at NPS was the last operational cathode from the Stanford FEL. When the FEL was decommissioned at Stanford, the gun was let up to

air and left at atmosphere for two years prior to being put back into operation at NPS. While the performance may have been degraded compared to when the cathode was first installed, it easily conditioned back to high voltages and successfully generated continuous beam on the first attempt. The vacuum requirements for thermionic cathodes are not as stringent as those for semiconductor cathodes which increases the thermionic cathode's robustness. Studies have shown thermionic cathodes can operate continuously for years without failure [5]. Although thermionic cathodes are not currently contemplated in the NPS superconducting, radio frequency (SRF) gun, the construction of the cavity does not preclude their use.

2. Metal and Semiconductor Cathodes

Metal and semiconductor cathodes operate using the photoelectric effect to liberate electrons from the cathode surface. Like the thermionic cathodes, these are immersed in an electric field to pull the electrons away from the surface as they are liberated. The electrodes or cavity shape are designed to pass the electron beam out of the gun and into the rest of the linear accelerator system. Unlike thermionic cathodes, a drive laser is required to provide the energy necessary to liberate the electrons.

In the photoelectric effect, a photon is absorbed by an electron. If the absorbed energy is sufficient to match or exceed the work function of the material, the electron can be liberated from the cathode surface. The work function is a property of the material and varies from $\sim 1 - 5$ eV. The work function also dictates the maximum laser wavelength required for a drive laser to be effective. For example, let us consider a niobium cathode that has a work function of 4.3 eV [6]; metal cathodes in general, have a work function of ~ 4 eV.

For a laser to liberate electrons from niobium, it must have a wavelength such that the photon energy is equal to or greater than the work function. The energy of a photon is given by $E = hc/\lambda$, thus, ignoring two photon effects, our drive laser must

have a wavelength shorter than

$$\lambda = \frac{hc}{4.3 \text{ eV}} = 288 \text{ nm} , \quad (\text{II.1})$$

which is in the ultraviolet range of the electromagnetic spectrum. Thus, using metal cathodes requires the use of short wavelength lasers. To use longer wavelength lasers, preferably in the visible range, it is necessary to move away from pure metal cathodes and consider semiconductor cathodes.

Metal cathodes are nearly as robust as thermionic cathodes, requiring similar handling and vacuum requirements. Oxide formation and possible cathode preparation requirements may make them less ideal for use. Their lifetime is long, but, in addition to the requirement for a UV drive laser, metal cathodes typically have poor quantum efficiency (QE), the number of electrons generated for each incident photon, often much less than 1%.

Usually, a semiconductor cathode is not a bare semiconductor. They are normally coated with a low work-function material from the alkali metals, such as cesium, which has a work function of 1.95 eV [6]. The addition of the alkali metal makes electron generation much better with QEs up to 10%. However, the lifetime of these types of cathodes is short, requiring reprocessing of the cathode to restore QE. Usually, the processing is done outside the gun to prevent putting low work function material into the gun cavity where reflected laser light from the cathode could cause electron emission where emission is not desirable. An additional benefit of the lowered work function is that a visible laser can be used as the drive laser, ensuring that electron emission only takes place on the cathode, not the electrode surrounding it. Long life, high QE cathodes remain an area of active research at multiple facilities.

Both types of cathodes are considered as candidates for the NPS SRF gun. The metal cathodes have already been demonstrated in both constant voltage (DC) and SRF systems. As the prevalence of alkali coated semiconductor cathodes in DC guns increases, there will be increased desire to try them in an SRF gun, especially if

QE lifetimes improve. At issue may be the higher vacuum requirement generally associated with this type of cathode, and migration of reactive metals from the cathode into the cavity.

3. Field Emission Cathodes

Field emission sources have been explored in other applications, but are just beginning to show promise as electron sources for linear accelerators. To date, no field emitter cathodes have been installed in a realistic gun environment, either DC or SRF. Of the three types of cathodes discussed, these are the least mature.

The cathode of greatest interest with respect to the NPS SRF gun is the diamond field emitter array (DFEA) being developed at Vanderbilt University. DFEAs are arrays of pyramidal structures that generate high current electron beams from their tips, shown in Figure 3. The thermal conductivity, mechanical strength, and potentially negative electron affinity of diamond make it ideal for this application [7]. Diamond field emitter tips have been seen to generate $100 \mu\text{A}$ per tip with 32 MV/m field on the tip (turn on is about 15 MV/m) [7]. Data for these devices comes from a lab setting in which low voltages of about 1 kV are applied across small gaps (on the order of μm) to generate the required turn-on fields.

Anecdotal evidence from conferences indicate that these DFEA are extremely robust, having been left in drawers unprotected for up to six months and returning to full current when put back under field. If this can be quantified, it would bode well for many types of electron guns. When coupled with diamond's resistance to contamination and oxide formation, the DFEA probably has the potential to be the most robust cathode candidate considered.

The field emitting cathode is also an interesting consideration for RF guns, as the only mechanism necessary to generate an electron beam is an applied field of sufficient strength. While the general RF fields will be strong enough for beam generation, the peak of a sinusoidal RF wave is not necessarily ideal for transport. However, frequency mixing schemes have been proposed for generating fast turn on,

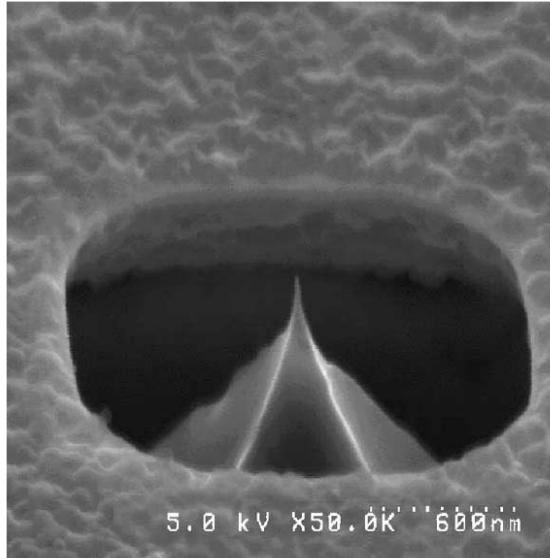


Figure 3. A scanning electron micrograph of a gated diamond field emitter. The pyramidal structure in the middle is the emitter with most of the electron emission being localized to the extremely sharp tip.

short duration pulses from field emitter cathodes using RF guns [8]. DFEA cathodes are being considered as part of the NPS SRF gun experimental plan.

B. DESIGN OF A UV DRIVE LASER SYSTEM

As part of the establishment of the NPS FEL group's laboratory spaces, it was necessary to develop a laser and transport system to conduct initial photoswitching experiments with the former Stanford DC gun. Additionally, the laser is to be used for separate test-stand experiments in the exploration of cathode performance in guns and gun-like structures. It is desirable to operate at multiple wavelengths to discriminate between emission materials and mechanisms where practicable. The system should also be transportable with minimal setup required before starting operation.

1. Minimalist Approach

To minimize starting costs and development time, the initial drive laser system was developed using a minimalist approach. The drive laser chosen was the Continuum Minilite II laser. It is a Nd:YAG (neodymium doped yttrium-aluminum-garnet) that operates primarily at 1064 nm with nonlinear crystals to convert the wavelength to 532 nm, 355 nm or 266 nm at 50 mJ, 25 mJ, 8 mJ, and 4 mJ, respectively. The pulse length is about 4-5 ns with a low, 1 - 15 Hz repetition rate, commensurate with our current capabilities to collect data. It is certainly not the drive laser for a high-power FEL system, but it is more than sufficient for the current experimental needs.

The initial beam line was designed to be attached to the Stanford DC gun and operate only at 266 nm to ensure photoemission from the metallic thermionic cathode surface. The optical path, shown in Figure 4, meets the bare minimum diagnostics and controls to be useful experimentally. There is one steering mirror for controlling the location of the laser spot on the cathode. The virtual cathode is set at the same distance from the steering mirror as the cathode, allowing the operator to track the laser spot size and motion via a video camera. Focusing of the laser spot is handled manually by inserting a focusing optic in a flip-in mount. For quantum efficiency measurements, an energy meter is used to measure the average pulse energy.

2. Optimized Design

As experience with the minimalist design was accumulated, the details developed for the additional diagnostics that would help in quantifying the laser beam as compared to the electron beam that was emitted. Prior to commencing further, more detailed cathode studies and a design for an upgraded laser table to fit in the same footprint was developed.

In the original design, to change the operating wavelength of the laser, it was necessary to open the laser head and change or remove crystals. The new table puts the doubling crystals on the table and intercepts the green laser light not upconverted

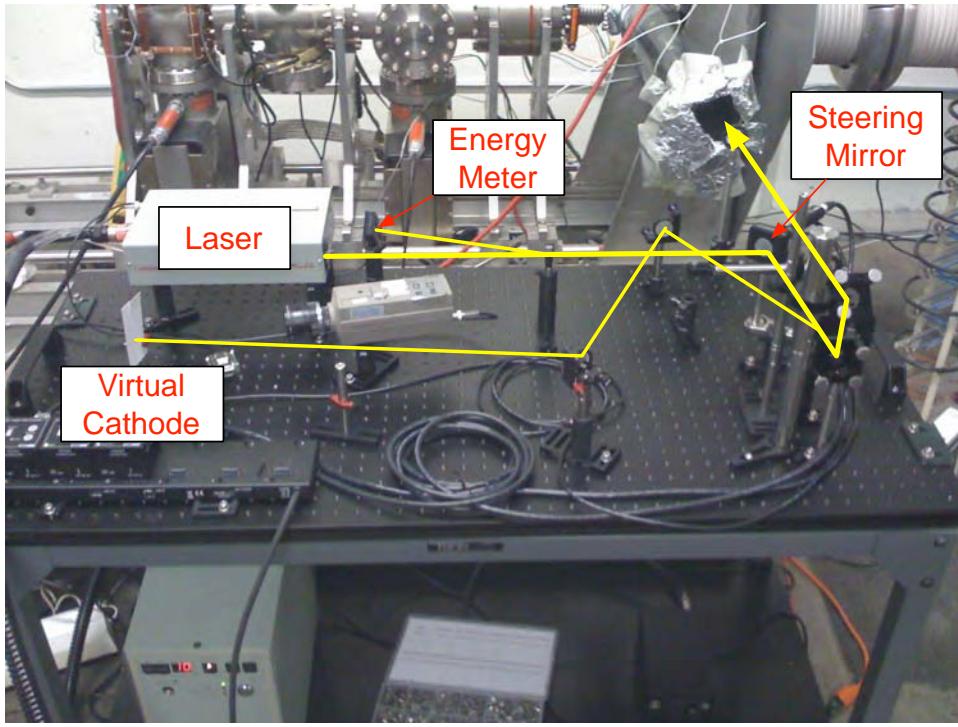


Figure 4. A minimalist drive laser design consisting of the bare minimum diagnostics and controls for performing QE experiments with the Stanford DC gun. The laser path is shown in yellow, terminating at the entrance to the vacuum system.

to UV passing through the dichroic mirror to use as a second color laser. This allows for the green and UV lasers to be coaligned for easier diagnosing of the laser spot position inside the test stands and for two-color experiments.

As the number of different experiments grew, the need to better quantify the laser pulse became more apparent. To add to the diagnostics of the laser beam, a photodiode for timing and temporal pulse shape has been installed. A beam profiler for measuring the transverse laser beam profile has also been installed. The need to frequently adjust beam spot sizes drove the decision to include a telescope for each beam line, allowing for spot size (diameter) adjustment from 4 mm to 50 μm . Rotatable polarizers were installed, allowing us to adjust the output beam energy of

each beam line individually from 0% to 100%. The final beam transport line, shown in Figure 5, is much more complicated, but also more agile and capable for future experiments.

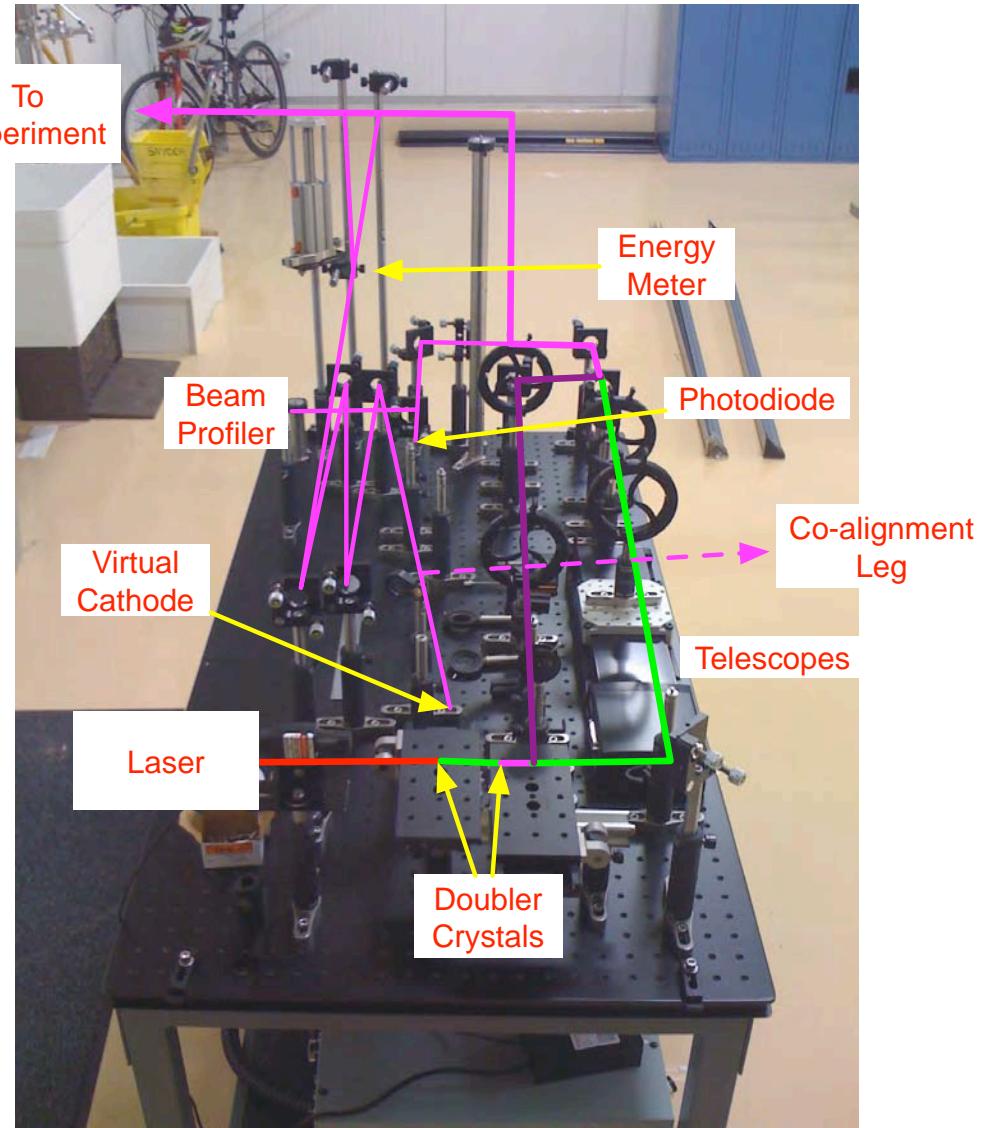


Figure 5. The updated NPS drive laser table showing the incorporation of two-color operation and the addition of more diagnostics and controls to better characterize and control the laser beam. In this photo, not all items have been mounted to the table.

III. SRF GUN/BOOSTER

Regardless of the electron source, we desire it to be located in a region of high accelerating field in order to bring the beam to high energy quickly. The longer the beam is left at lower energies, the greater the impact of space charge on the beam, increasing its emittance and making transport more difficult. This accelerating gradient can be applied in one of two ways—through a constant voltage or through the application of a periodic varying voltage.

Constant voltage, or direct current (DC), electron guns are well developed and used in many electron sources today (e.g., [9], [10] and [11]). In these systems, a constant voltage is established between the cathode (usually at negative potential) and the anode (usually at ground) and possibly some intermediate electrodes for electric field shaping. Figure 6 shows a drawing of the former Stanford Superconducting Accelerator electron gun as an example. Modern versions of these guns are typically designed to a standoff voltage of between 300 kV and 750 kV [9], [10]. They achieve fairly high fields at the cathode face, but the overall gradients are typically much lower than can be achieved with a varying accelerating field as applied by a radio-frequency (RF) field.

The RF approach has a significant advantage over the DC case, as the time-averaged fields are much lower than the peak fields. As a result, the overall gradient can be increased, greatly reducing the time the electron beam is at low energy. Unfortunately, these improvements come with some significant costs. The beam now does not see a constant accelerating gradient, so timing of the injection becomes critical. Instead of a DC power supply, RF power supplies must be used, and waveguides must be used to port the RF power into a resonant cavity. The cavity itself must be tuned and designed to a specific frequency. The fact that the accelerating voltages in an RF cavity oscillate mean that there are only certain times when the fields are conducive

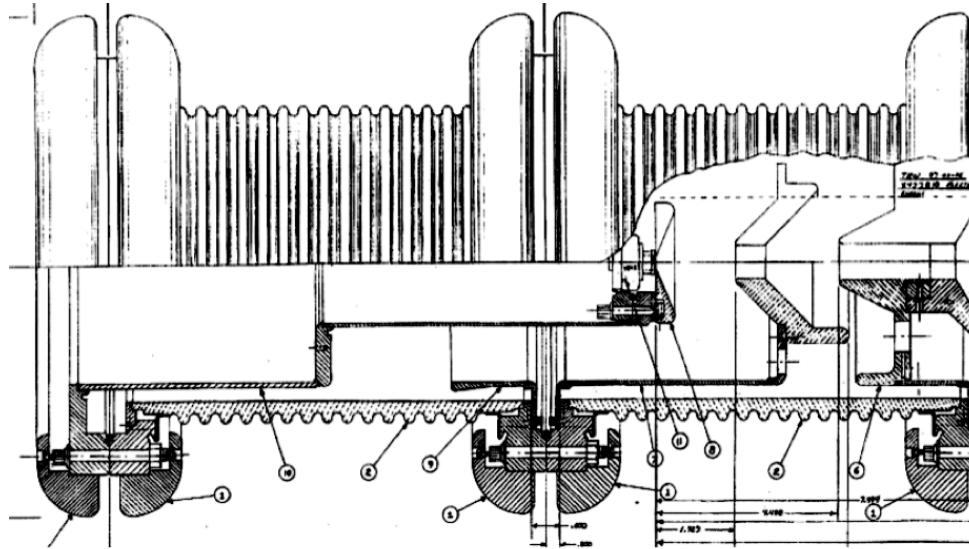


Figure 6. This is a drawing of the former Stanford Superconducting Accelerator DC gun from files delivered with the gun. This gun operated with a thermionic cathode that was gated with an applied grid voltage. It routinely operated at voltages > 200 kV and had a maximum gradient (with cathode at -240 kV and first electrode at ground) of 2.3 MV/m. It typically operated at some lower voltage, allowing the first electrode to provide beam focusing.

to acceleration of charged particles, leading to pulsed beams. Higher gradients also necessitate higher wall currents, leading to heating of the cavity and parasitic power loss.

In the RF accelerating scheme, there are two choices that can be explored in terms of cavity material type—normal conducting and superconducting. Both types of conductors provide the structure for resonant cavities; however, the power dissipated in normal conducting structures is substantially higher than that dissipated in superconducting ones [12], due to increased ohmic heating in normal conducting cavities. While there are additional engineering and refrigerator requirements for superconducting cavities, they are more than compensated for by decreased RF power demands, especially in continuous wave applications. Table 1 shows the power comparison between copper and superconducting cavities for modest accelerating gra-

dients. We see immediately that, even for moderate gradients, the superconducting cavity provides us with significant power savings.

Table 1. AC power requirements for superconducting and normal conducting cavities at 1 MV/m and 5 MV/m. Q_0 is the quality factor of the cavity, r_a/Q_0 is the geometric shunt impedance, and E_{acc} is the accelerating gradient. From [12].

Option	Super-conducting	Normal Conducting
Q_0	2×10^9	2×10^4
r_a/Q_0 [Ohm/m], RF frequency = 500 MHz	330	900
P/L [Watt/m] for $E_{acc} = 1$ MV/m	1.5	56,000
AC Power [kW/m] for $E_{acc} = 1$ MV/m	0.54	112
AC Power [kW/m] for $E_{acc} = 5$ MV/m	13.5	2,800

As of this writing, there is only one superconducting RF electron gun in use as an FEL electron source in the world. The superconducting RF electron gun used at the Forshungszentrum Dresden-Rossendorf in Rossendorf, Germany, shown in Figure 7, is the electron source for the Free Electron Laser (FEL) at the Electron Linear accelerator with high Brilliance and low Emittance (ELBE). The prototype (a single half-cell) first developed beam in 2003 [13] and led to the production model gun, a 3.5-cell version, which made first beam in late 2007 [14]. As this gun (referred to from this point as the “Rossendorf gun”) is the only superconducting RF (SRF) gun currently in use, its parameters are of particular interest for comparison purposes. Design simulation results of a similar geometry 650 MHz Rossendorf gun are shown in Table 2, which is a more relevant frequency for comparison.

The Naval Postgraduate School (NPS) has established an FEL experimental program to complement its robust theoretical program, and desires to explore innovative accelerator, beamline, and undulator designs to further FEL technology [17], [18]. To this end, it was proposed to develop and build a superconducting electron gun for use in our lab. As a brief overview, the gun is designed to operate at 500 MHz for frequency compatibility with the main linear accelerator. The gun should be a

Table 2. Characteristic design parameters of a 650 MHz Rossendorf SRF electron gun, from simulation. Parameters are taken from an unpublished report provided by Dr. J.W. Lewellen. The symbols for the parameters are defined in the list of symbols.

Parameter	Value
f_0	650 MHz
q_b	1 nC
σ_t	24.6 ps, FWHM
E_0	26.48 MV/m
ϵ_n	1.19 mm-mrad
σ_x^*	0.97 mm
t_b^*	9.3 ps
KE	13.8 MeV
δ	0.23%

* At the end of the simulated beamline, $z = 1.42$ m (from the cathode surface)

compact, superconducting RF gun with high gradient and good beam quality. Having accomplished the design, development and construction of an initial prototype, NPS has the first SRF electron gun in the United States and is undergoing testing. The gun design, manufacture, and testing will be discussed extensively in the remainder of this chapter.

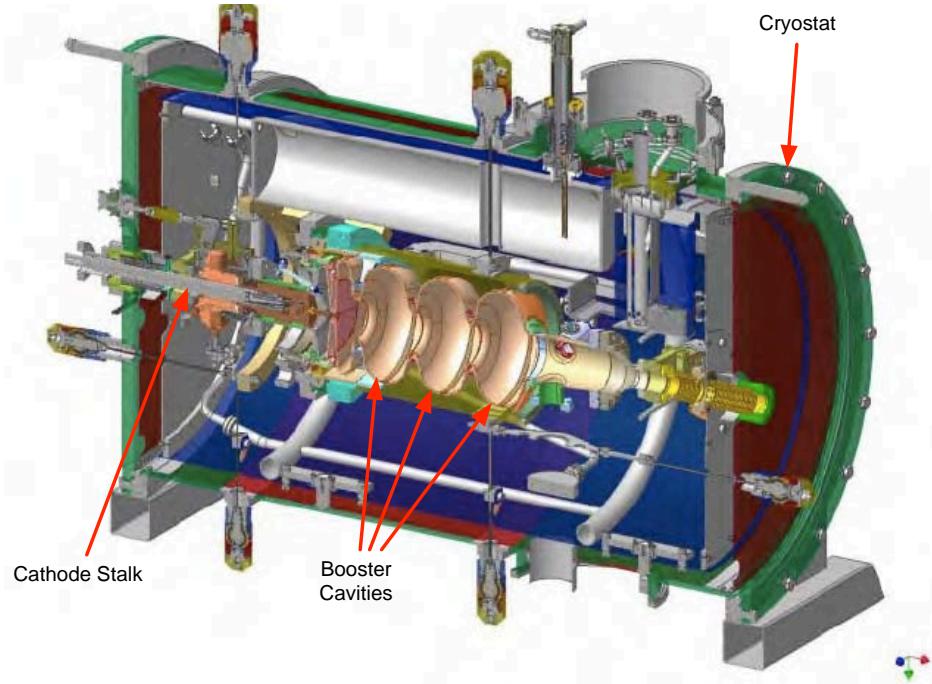


Figure 7. The Rossendorf gun is a superconducting photoelectron gun consisting of 3.5 elliptical cells, operating at 2 K and 1.3 GHz [15]. The diagram also shows all the pertinent components for an SRF gun, including the cryostat. The cathode stalk is inserted from the left. For length scale, the 3.5 cell structure is approximately 40 cm from cathode surface to the end of the last cell, yielding a cryomodule approximately 1.5 m in length and 80 cm in diameter. From [16].

A. RF CAVITY TYPES

There are many different types of cavities that can be used in RF accelerating structures. The cavity shape determines the accelerating field shape the electrons will experience in the gun. Changing the RF frequency, or introducing a second frequency into a cavity can also alter the cavity's behavior. The ideal cavity mode for an electron gun is one in which there is a large axial electric field with small radial electric fields at an easily obtained RF frequency. The axial electric field requirement can be met by designing a cavity for a transverse magnetic mode (the TM_{010} mode, for example). The oscillatory nature of the RF fields requires that any charged particle beam will be pulsed in order to properly match with an accelerating field in the cavity.

The first RF cavities were based upon the simple pillbox structure. The pillbox is a simple zeroth order cavity shape, allowing for simplicity in manufacture and modeling. This structure provides the jumping-off point for the much more advanced structures that are in use today.

1. Right Cylinder Pillbox Cavity

As mentioned, the simplest RF cavity shape is the pillbox structure. This is a right cylindrical structure with a cut-out in the end caps for beam entrance and exit as shown in Figure 8. The magnetic field has the desired transverse behavior and the electric field lies along the long axis of the cavity with the strongest fields on axis.

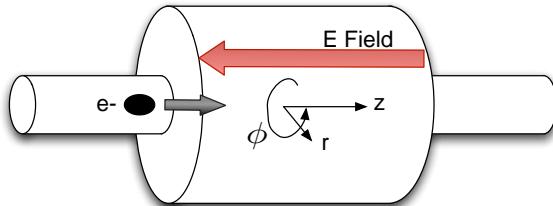


Figure 8. A generalized pillbox cavity showing the transit direction of a particle beam.

As the pillbox cavity's fields can be derived analytically, we include some of the results as they are illustrative for considering some of the more advanced cavity shapes. We follow the general process for solving for waveguide modes as can be found in several electromagnetism textbooks, [19] and [20], for example. We assume a plane wave of the form

$$\mathbf{E}(\mathbf{x}, t) = \tilde{\mathbf{E}}(r, \phi) e^{i(kz - \omega t)}, \quad (\text{III.1})$$

where r is the radial distance from the propagation ($\hat{\mathbf{z}}$) axis, ϕ is the angular distance from the $x - z$ plane, k is the wavenumber, and ω is the angular frequency. We also desire to separate the transverse effects from the longitudinal effects, so we define an operator, $\nabla_{\perp} \equiv \langle \partial/\partial x, \partial/\partial y, 0 \rangle$, which captures the transverse effects such that

$$\nabla^2 \equiv \nabla_{\perp}^2 + \frac{\partial^2}{\partial z^2}. \quad (\text{III.2})$$

We apply the wave equation to this field,

$$\begin{aligned} \left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} &= 0 \\ \left(\nabla_{\perp}^2 + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} &= 0 \\ \left(\nabla_{\perp}^2 + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \tilde{\mathbf{E}} e^{i(kz - \omega t)} &= 0 \\ \left[\nabla_{\perp}^2 + \left(\frac{\omega^2}{c^2} - k^2 \right) \right] \tilde{\mathbf{E}} &= 0. \end{aligned} \quad (\text{III.3})$$

Since the wave equation applies to both the \mathbf{E} and \mathbf{H} fields and we can define them in a similar fashion (for a plane wave propagating in the $\hat{\mathbf{z}}$ direction), then $\tilde{\mathbf{E}}$ in equation (III.3) can be replaced with $\tilde{\mathbf{H}}$ to describe the magnetic field.

Since our expected cavity material can be considered a nearly perfect conductor, our boundary conditions are the usual boundary conditions for a waveguide,

$$E_z|_s = 0 \text{ and } H_{\perp}|_s = 0, \quad (\text{III.4})$$

where $|_s$ denotes "on the surface."

Confining our discussion to the fundamental mode, and assuming we intend to use the cavity to accelerate beams, then the solutions we desire must have a non-zero E_z . This confines our solutions to the transverse magnetic (TM) modes. The E and H fields in TM modes are related through

$$\mathbf{H}_\perp = \pm \frac{\epsilon_0 \omega}{k} (\hat{\mathbf{z}} \times \mathbf{E}_\perp) . \quad (\text{III.5})$$

The transverse electric fields are related to the accelerating field by

$$\mathbf{E}_\perp = \pm \frac{ik}{\Gamma^2} \nabla_\perp E_z \quad (\text{III.6})$$

where $\Gamma^2 \equiv (\omega^2/c^2) - k^2$ and E_z satisfies equation (III.3). These are traveling wave solutions, but we desire standing wave solutions to be applicable to a cavity [12].

To make our solution applicable to a cavity instead of a waveguide, we impose the boundary conditions necessary at the end caps, $E_\perp = 0$ at $z = 0, a$. This constrains the z components, allowing us to find the accelerating field,

$$E_z(\mathbf{x}, t) = \tilde{E}(r, \phi) \cos\left(\frac{p\pi z}{a}\right) e^{-i\omega t} , \quad (\text{III.7})$$

where the boundary conditions imply only certain wavelengths can satisfy the wave equation and be present in the cavity ($k = p\pi/a$) and $p \in \{\mathbb{J} \geq 0\}$, where \mathbb{J} is the set of integers. Thus

$$\Gamma_{mnp}^2 = \frac{\omega_{mn}^2}{c^2} - \left(\frac{p\pi}{a}\right)^2 , \quad (\text{III.8})$$

where $n \in \{\mathbb{J} \geq 0\}$. Putting these together into equation (III.6), it can be shown that

$$\mathbf{E}_\perp = \frac{-m\pi}{a\Gamma_{mnp}^2} \sin\left(\frac{p\pi z}{a}\right) \nabla_\perp \tilde{E}(r, \phi) e^{-i\omega t} . \quad (\text{III.9})$$

For TM modes, we know $H_z = 0$. However, we still need to find H_\perp . Using equation (III.5) and our \mathbf{E} fields, we find

$$\mathbf{H}_\perp = \frac{i\omega_{mn}}{\eta c \Gamma_{mnp}^2} \cos\left(\frac{p\pi z}{a}\right) \hat{\mathbf{z}} \times \nabla_\perp \tilde{E}(r, \phi) , \quad (\text{III.10})$$

where η is the impedance of free space ($\sqrt{\mu_0/\epsilon_0}$).

Finally, stipulating that our pillbox cavity has radius R , then the lowest order mode (TM_{010}) for our pillbox cavity has field components

$$\begin{aligned} E_r, E_\phi, H_r, H_z &= 0 \\ E_z &= E_0 J_0 \left(\frac{2.405r}{R} \right) e^{-i\omega t} \\ H_\phi &= -i \frac{E_0}{\eta} J_1 \left(\frac{2.405r}{R} \right) e^{-i\omega t}, \end{aligned}$$

where J_n is the n -th order Bessel function. The 2.405 comes from the first zero of the Bessel function.

In general, the radial electric fields can be related to the longitudinal electric field by using $\nabla \cdot \mathbf{E} = 0$, to see how the radial fields behave as E_z changes along the cavity. This is true also for more complicated cavity shapes. We assume here that the fields are axisymmetric ($E_\phi = 0$),

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 0 \\ \frac{1}{r} E_r + \frac{\partial E_r}{\partial r} + \frac{1}{r} \frac{\partial E_\phi}{\partial \phi} + \frac{\partial E_z}{\partial z} &= 0 \\ \frac{1}{r} E_r + \frac{\partial E_r}{\partial r} + \frac{\partial E_z}{\partial z} &= 0. \end{aligned} \tag{III.11}$$

We take the derivative of equation (III.11) with respect to r ,

$$\begin{aligned} \frac{\partial}{\partial r} \left[\frac{1}{r} E_r + \frac{\partial E_r}{\partial r} + \frac{\partial E_z}{\partial z} \right] &= 0 \\ -\frac{1}{r^2} E_r + \frac{1}{r} \frac{\partial E_r}{\partial r} + \frac{\partial^2 E_r}{\partial r^2} + \frac{\partial^2 E_z}{\partial r \partial z} &= 0 \end{aligned}$$

ignoring terms involving second derivatives because we assume we are near the electric axis of the cavity where the curvature should be small,

$$\begin{aligned} \frac{1}{r} \left(\frac{\partial E_r}{\partial r} - \frac{E_r}{r} \right) &\approx 0 \\ \frac{\partial E_r}{\partial r} &\approx \frac{E_r}{r}. \end{aligned} \tag{III.12}$$

Substituting equation (III.12) into equation (III.11), we find

$$\begin{aligned} \frac{1}{r}E_r + \frac{1}{r}E_r + \frac{\partial E_z}{\partial z} &\approx 0 \\ E_r &\approx \frac{-r}{2} \frac{\partial E_z}{\partial z}, \end{aligned} \quad (\text{III.13})$$

which can be extremely useful in estimating the radial fields for cavities for which the axial fields are known.

2. Tesla Cavity

The first historical modifications to the pillbox structure for SRF applications were to round the corners where the end caps meet the longitudinal structure to improve surface current conduction by removing the 90-degree turn (and weld) and to improve the ability to clean the cavities. Removing surface contaminants is of utmost importance in superconducting cavity construction, as will be discussed later in the subsection on cavity cleaning and assembly. Cavity design modifications continued to improve manufacturability and performance.

Probably the largest driver in early alteration of cavity cell designs is multipacting, an effect where field-emitted electrons are ejected into the cavity fields and then impact the cavity wall very near their origin, liberating secondary electrons that then are ejected into similar paths cascading into large populations that pull energy from the cavity fields. Once enough electrons are liberated, the cavity's accelerating gradient is limited. Figure 9 shows a trajectory trace from a multipacting simulation. The first successful attempt to limit multipacting was to curve the corners of the pillbox, as in the HEPL (High Energy Physics Laboratory) accelerator at Stanford University [21]. These cavities can be seen in Figure 10.

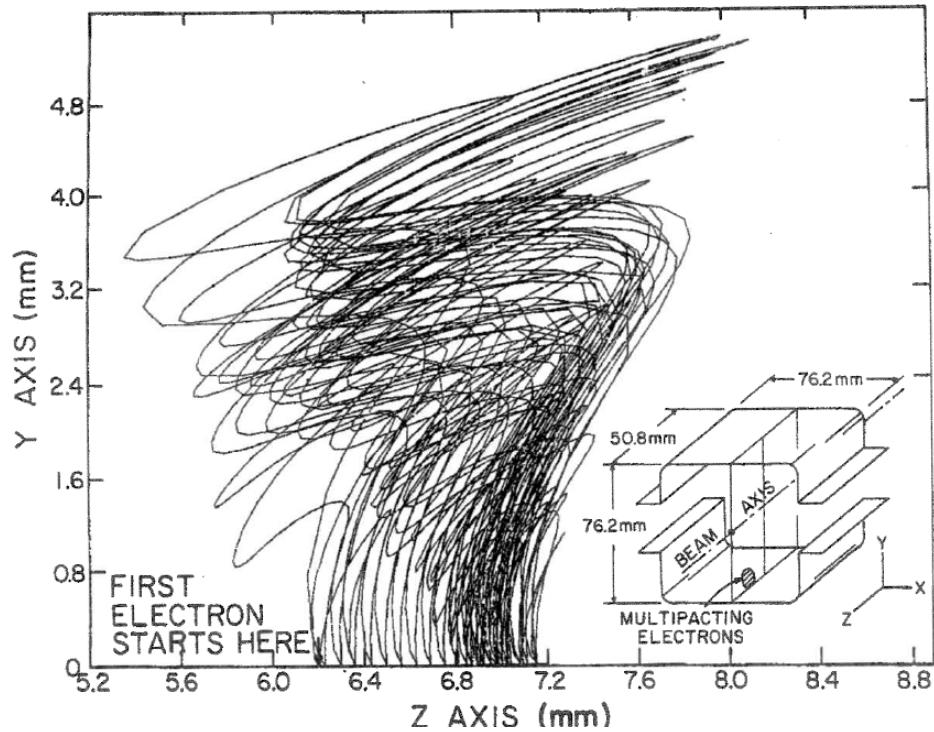


Figure 9. Single point multipacting trajectories in a muffin-tin cavity [22]. This diagram demonstrates how a single electron can, over multiple RF cycles impact very near its origin causing localized heating, and for sufficient impact energy, liberate a significant population of secondary electrons. These secondaries will follow very similar paths, compounding the heating and cavity power limiting effects.



Figure 10. A seven cell HEPL structure. From [12].

Accelerator cavities continued to evolve, eventually being modified to an elliptical shape, as seen in Figure 11, allowing the cavity to be manufactured from two cell halves, with only one equatorial weld. Additionally, the continuously curved cavity walls also significantly reduce multipacting as the liberated electrons are forced to migrate to the cell equator, as shown in Figure 12, where the electric field nearly vanishes. The Tesla cavity geometry, developed for the TESLA linear collider, has become the predominant cavity type used in most modern superconducting electron accelerators, usually operating at frequencies above 1 GHz and gradients at or above 10 MV/m.

3. Quarter-Wave Cavity

The cavities that have been discussed to this point have all had a hidden assumption built into their structures. That assumption is that the cavities are designed to accelerate speed-of-light particles ($\beta \simeq 1$), where $\beta \equiv v/c$. The structures to accelerate low β particles can have exotic shapes to provide for optimal acceleration. Of interest to us is the so-called quarter-wave structure, an example of which is shown in Figure 13. The components of the quarter-wave structure are a pillbox type outer skin and a coaxial line that intrudes into the interior of the cavity.

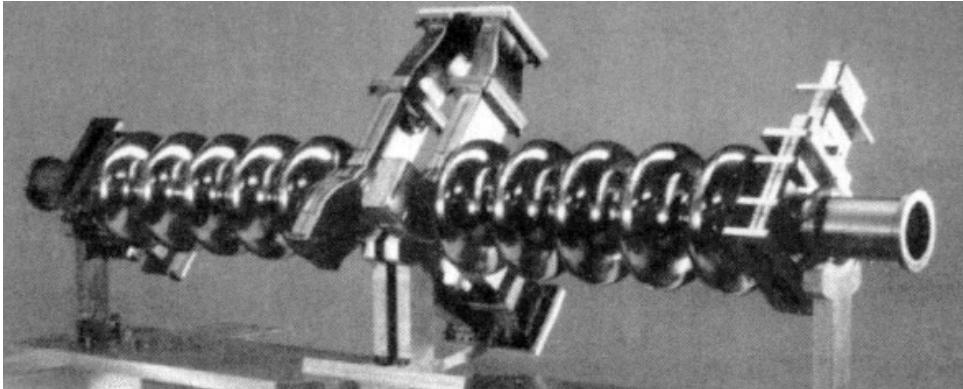


Figure 11. A pair of 1500-MHz 5-cell, niobium cavities. The structures in the middle are RF waveguides. From [12].

As seen in Figure 13, this particular cavity is designed so that a beam is passed not through the end caps of the cavity, but through the side of the cavity and the end of the coaxial line. The coaxial structure allows for a TEM mode to be established along its length, providing an alternating accelerating gradient (a radial electric field in the cavity frame) across the two gaps the beam would see. While the accelerating gradient is reversing, the beam is shielded in a drift tube until the gradient again reverses and provides a second acceleration as the beam exits the cavity. Driving frequencies in quarter-wave structures are typically much lower than in $\beta = 1$ cavities, which decreases the thermal load on the cavity due to the alternating currents induced in the cavity walls.

We propose to use the quarter-wave structure for the acceleration of electrons and the development of a compact electron beam source. Our usage of the structure will not be in its usual mode; instead we will establish the accelerating gap between the end of the coaxial line and an exit port through the opposite endcap. This gives us a cavity that should be easy to manufacture, will operate at a frequency low enough

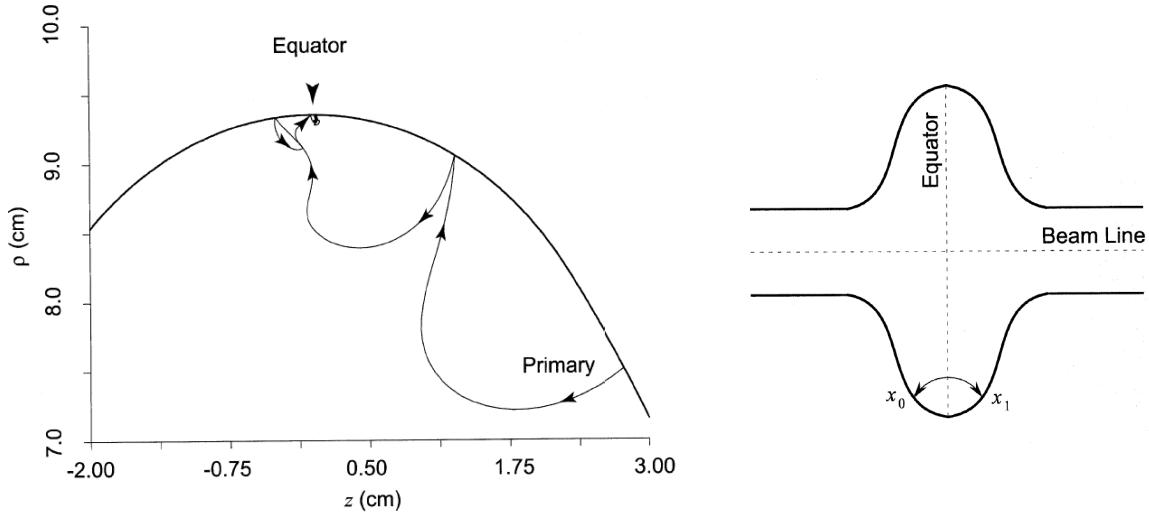


Figure 12. The typical multipacting trajectory in an elliptical cavity is shown on the left. On the right is a different form of resonant multipacting that is not eliminated by the use of the elliptical shape. After [12].

to operate at 4.2 K, and can be used as both a booster and, with the addition of a cathode stalk, as an electron gun. The remainder of this chapter will address the major components' manufacture, modeling, and testing.

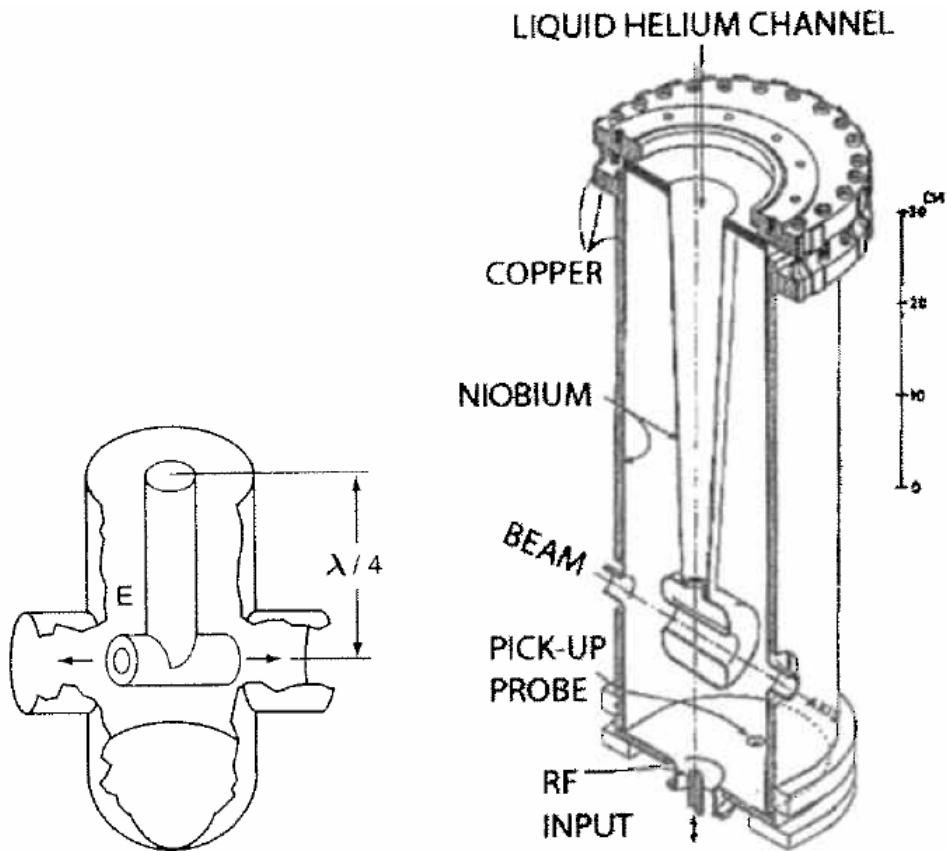


Figure 13. The diagram on the left shows the generalized quarter-wave structure. On the right is a drawing of a quarter-wave structure used for the acceleration of heavy ions. The central coaxial line in this device is filled with liquid helium for cooling. After [12].

B. FABRICATION

The NPS quarter-wave, superconducting cavity was initially conceived as a booster cavity to be used in conjunction with the former Stanford DC electron gun. Together, the two devices would make up the injector section of the NPS FEL, providing electrons nominally at 1.5 MeV (\sim 250 keV from the gun and 1.2 MeV from the booster). The booster cavity is designed to handle significant bunch charge (1 nC) at fairly low repetition rate (initially 10s of Hz building to kHz as the experiments progress) to minimize the liquid helium cooling and RF power required. To also decrease the cryogenic cooling and auxiliary support equipment necessary, the cavity is designed to be operated at 4.2 K, which is above the superfluid helium temperature of 2 K, but sufficiently cold to cause niobium to become superconducting ($T_c = 9.2$ K). The fundamental frequency was chosen to be 500 MHz in order to be compatible with the linear accelerator to which it is to connect. In this case, the accelerator operates at 1.3 GHz, meaning that this system's highest possible repetition rate (limited by when the various frequency RF fields throughout the beamline are properly in phase) is 100 MHz. Since a bunch can only pass through the entire line (due to the required phase alignment) once every 13 cycles, this limits the maximum duty factor of the accelerator sections to about 8%, reducing RF and cooling power. As the NPS FEL is designed to be a low duty factor system (on the order of Hz to kHz for the foreseeable future), the 100 MHz bunch frequency does not place a significant constraint on system operation. For reference in the remainder of this section, the final NPS 500 MHz superconducting gun/booster schematic is shown in Figure 14. In this layout, electrons are injected from the left and are accelerated as they pass through the nosecone and out the cavity to the right.

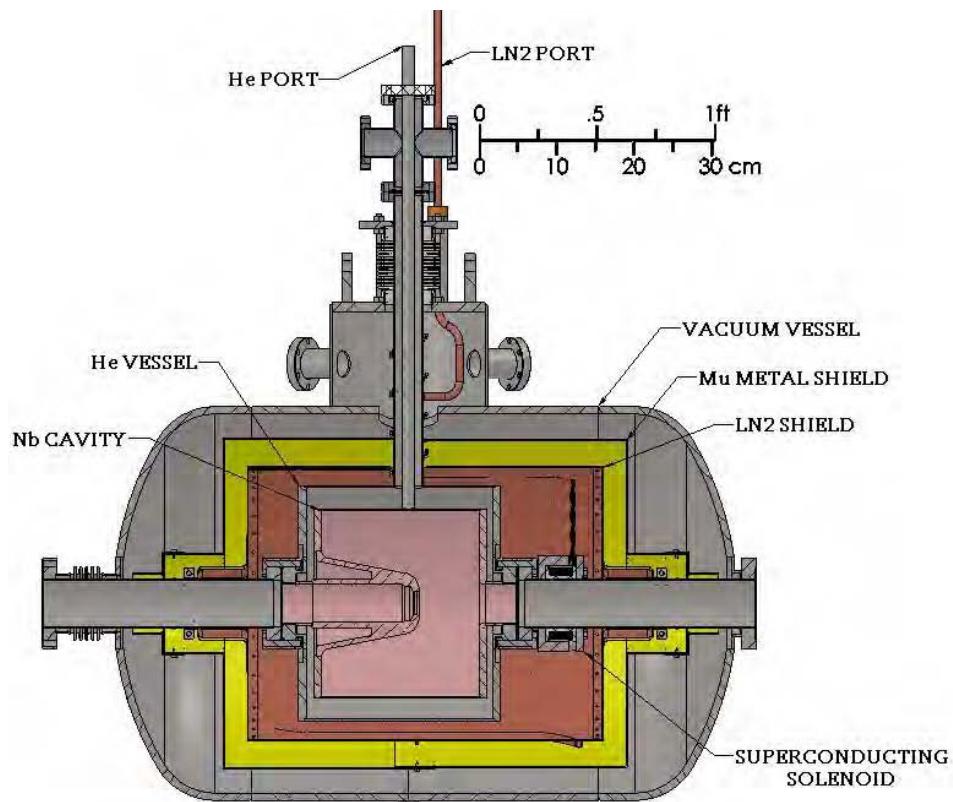


Figure 14. The NPS 500 MHz superconducting gun/booster as built by Niowave, Inc. Shown in the drawing are the ports for flowing liquid nitrogen and helium as well as the various enclosures inside the cryostat. Not shown are the locations of various internal temperature sensors, the cathode stalk assembly, or coupler assembly.

1. Cavity

Initial cavity design work was performed using the Superfish¹ family of codes. The cavity structure was optimized to obtain the desired fundamental frequency and field strength by altering the cavity dimensions, including coaxial line diameter, overall cavity diameter, cavity length, and coaxial line length. To explore how the coaxial parameters affect the fundamental frequency, a right-cylindrical pillbox cavity with the same length and radial dimensions of the intended NPS gun, as shown in Figure 15, is examined as a right-cylindrical coaxial line was inserted from 0% to 95% of the cavity length. Figure 16 illustrates how these parameters affect the cavity's fundamental frequency. As can be seen in Figure 16, the coaxial line radius has significant impact on the cavity frequency, depressing it until the coaxial line radius reaches half of the cavity radius. Increasing the coaxial line radius beyond 50% of the cavity radius only serves to drive the fundamental frequency back up. Insertion of the coaxial line further into the cavity depresses the cavity frequency almost linearly once inserted beyond 25% of the cavity length. By varying both parameters, a desired frequency and accelerating gradient and gap can be designed for a particular cavity.

While performing this design exercise for the production cavity, we also ensured that the fields expected in the cavity did not exceed fundamental limits on the maximum surface magnetic field. For niobium, the peak surface magnetic field must be less than the superheating field, 0.23 T (183 kA/m). This H_{pk} limits the peak electric field that can be obtained in the cavity to approximately 55 MV/m as well [12]. Based on the quarter-wave structure, we expect that the peak magnetic field will occur in the coaxial line section of the cavity while the maximum surface electric field (which determines the amount of field emission we expect) should occur along the nose cone.

After optimization, we find via Superfish field maps of the booster configuration that, for the nominal 1.2 MV gap voltage, our maximum magnetic field is on

¹The Superfish family of simulation codes will be discussed in Chapter IV.A.1.a

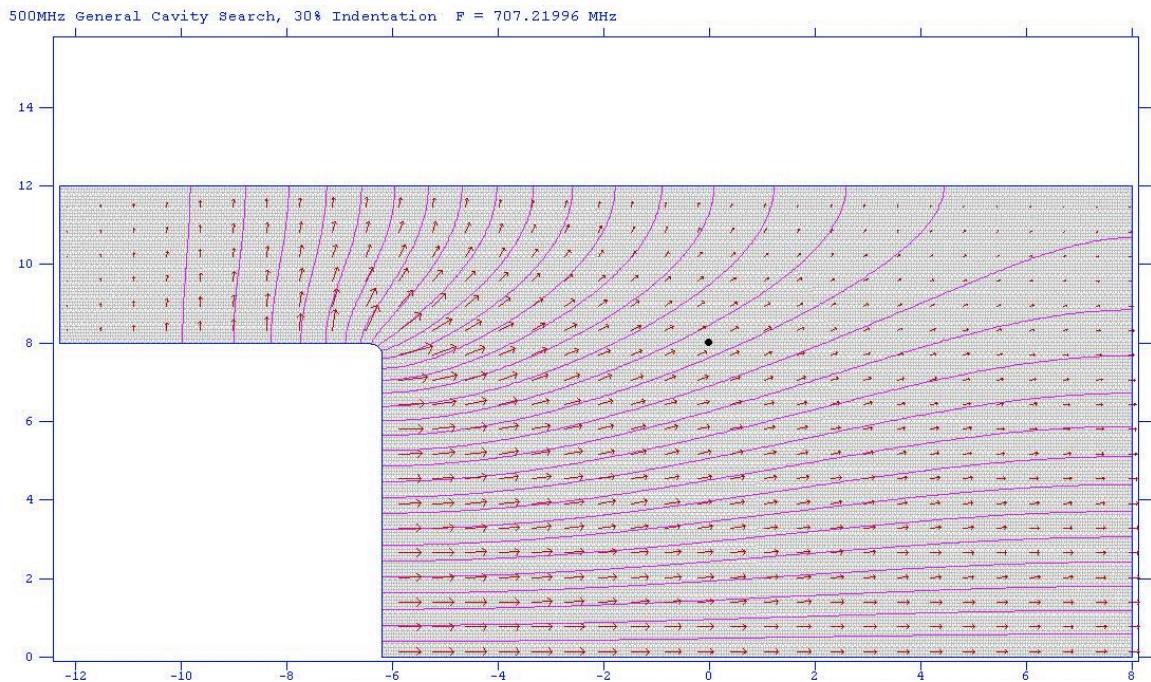


Figure 15. An example pillbox used to explore fundamental frequency shifts due to insertion of a coaxial line. Coaxial lines are inserted from the right and insertion lengths and coaxial line radius are varied. The pillbox structure is 20.305 cm long and has radius 12.000 cm. Pink lines indicate electric field lines and red arrows show relative electric field strength.

the order of 65 mT occurring approximately two-thirds of the way back from the nose cone along the coaxial line, while the maximum electric field of 49 MV/m occurs just inside the tip of the nose cone. Also of interest is the field that will be experienced at the cathode surface in the gun configuration. Since the cathode stalk can be adjusted to the compression and extension limits of the bellows, we can adjust the cathode's surface from a retracted position 14.23 mm behind the nose cone to about 1.51 mm beyond the nose cone. Since extending the cathode into the accelerating gap results in the emitted electrons experiencing significant defocusing fields upon emission, we limit our consideration of positions to those from insertion flush with the nose cone to retraction still sufficient for field emission cathodes ($|E| > 10-15$ MV/m) as

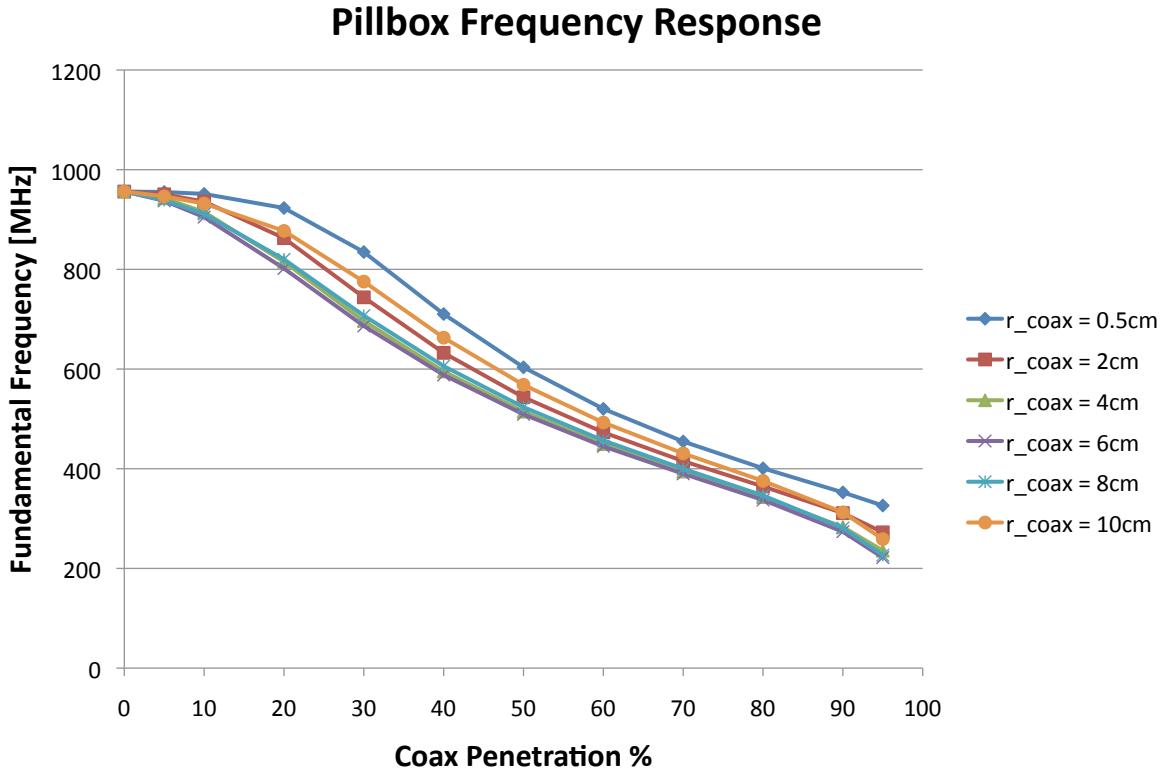


Figure 16. Fundamental frequency effects in a pillbox cavity seen when varying the coaxial line radius and penetration depth (expressed as a percentage of the pillbox length) into a simple pillbox structure. The pillbox structure is 20.305 cm long and has radius 12.000 cm.

discussed in the previous chapter. Figures 17 and 18 show the accelerating field, E_z (on axis), and transverse field, E_r ($r = 1$ cm), for various cathode positions and the booster configuration.

We see from Figures 17 and 18 that there is an obvious trade-off between peak accelerating fields on the cathode surface and radial focusing as the beam is launched from the cathode surface. From an applications standpoint, assuming the cavity achieves its designed gradient, all cathode positions forward of 6.5 mm retraction would be usable for field emitter cathodes which require approximately 10 MV/m for turn-on [23], [7].

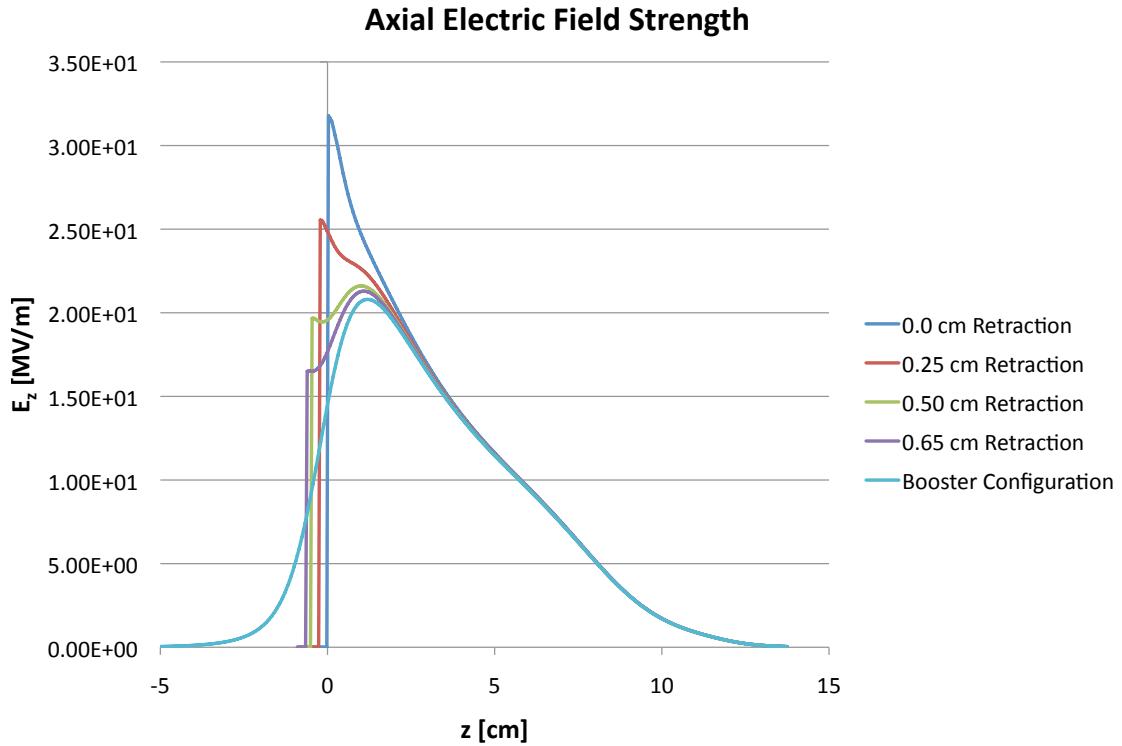


Figure 17. Plots of the axial electric field, E_z , along $(r, \theta, z) = (0, 0, z)$ as the cathode position is changed within the cavity. Retraction distances are measured relative to the cavity nose cone ($z = 0$). Data is from Superfish simulations.

Having a radial focusing field at the cathode surface is beneficial for many reasons. Most important is countering the space-charge effects tending to increase the beam's size and emittance until the beam can be accelerated significantly. Figure 18 provides radial field strength curves for various configurations of the NPS quarter-wave cavity. For cathode retractions less than 3.0 mm, there is no focusing as the beam leaves the cathode. Maintaining a retraction beyond this point provides focusing for the beam as it leaves the cathode until it can be accelerated, countering the effects of space charge, but at the expense of accelerating field, E_z . This indicates the possibility that this gun configuration could be useful for high bunch charges. Also interesting is that when used in the booster configuration, there are significant focusing fields (> 2 MV/m) as the beam passes through the cavity. In many injectors, acceleration after

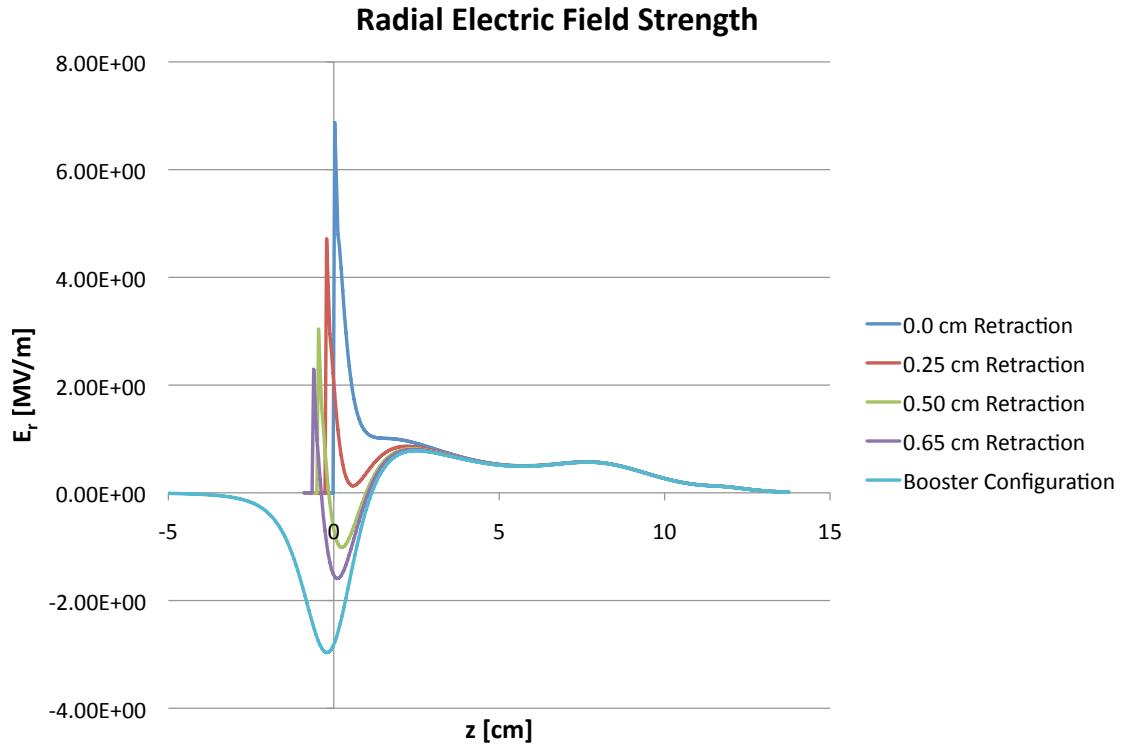


Figure 18. Plots of the radial electric field, E_r , along $(r, \theta, z) = (0.5 \text{ cm}, 0, z)$ as the cathode position is changed within the cavity. Retraction distances are measured relative to the cavity nose cone ($z = 0$). Negative E_r indicates a radially focusing field. Data is from Superfish simulations.

the gun is delayed until after a focusing element in order to preserve bunch charge prior to further acceleration. It may be possible to eliminate that focusing element and decrease the size of an injector using this cavity as a booster.

Actual manufacture of the cavity was performed at Niowave, Inc., facilities in Lansing, MI. Its facility includes almost all machines, materials, and facilities for manufacture and testing of accelerating structures. With the exception of electron beam welding of the cavity, all manufacture and processing of the NPS cavity was performed on-site at Niowave.

The cavity began as the four primary components shown in Figure 19. The nose cone cylinder and end plates were cut from a high residual resistivity ratio (RRR), large grain niobium ingot. The hollow cylinder was formed by bending small grain sheet niobium. Higher values of RRR material are desirable, as they result in decreased heat load generated through the high alternating currents in the cavity surfaces when operating and increased heat transport (thermal conductivity) of the generated heat. The ingot material used had a RRR value of 180, while the sheet material had a RRR value of 300. These are representative of “typical” values of commercially available niobium and no significant attempt was made to procure extremely pure niobium for this cavity.

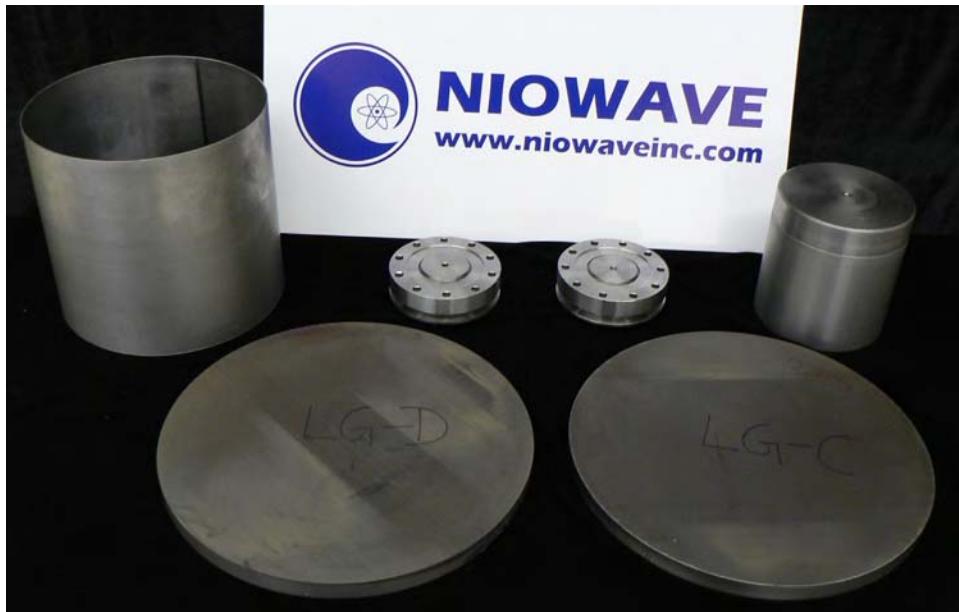


Figure 19. The component parts of the NPS 500 MHz quarter-wave cavity. The two large disks at the bottom are slices from a large grain niobium billet and form the end plates of the cavity. The cylinder on the left is rolled sheet material that makes up the outer conductor. The solid cylinder on the right is cut from the same billet as the end plates and makes up the inner conductor. The two remaining parts in the picture are niobium-titanium flanges that mate the cavity to the beamline.

The two end plates and nose cone were milled to final thickness using standard machining processes. The nose cone is hollowed between the inner beam pipe and outer coaxial surface to allow liquid helium to reach as close as possible to the high current areas of the cavity.

Once the cavity was assembled, the first set of verification experiments were performed. These experimental results are presented and discussed in the simulation and experiment chapter.²

2. Coupler

The primary function of the coupler is to transport RF energy from a source into the stored fields of the cavity. In the case of many cavities, this is accomplished either through waveguides or RF antennae. As our cavity is extremely compact and our design philosophy included minimizing the number of cavity penetrations, it was decided not to use a cavity penetrating RF antenna for the first prototype.

Our design choice was to use an axial waveguide coupler. This has the benefit of not requiring an additional cavity penetration as the output beam tube is used to convey the evanescent wave from the end of the coaxial waveguide to the cavity. Additionally, the hollow coaxial line allows for electron beam extraction. The coupler drawing is shown in Figure 20. The two primary drawbacks to this design are that the coupler decreases the available aperture for extracting the electron beam and the required drift space occupied by the coaxial coupler allows space charge time to expand the beam. Simulations described later indicate the resulting aperture decrease is not a major issue except at lower cavity gradients.³ However, the increased drift space required by this design and other factors may indicate that a different form factor may provide substantial benefits in this cavity's use as a booster or gun.

The configuration shown in Figure 20 has two external connections. The connection on the bottom is an RF connection for coupling RF power onto the coaxial

²See Chapter IV.

³See Chapters IV.A.2., IV.A.3. and IV.B.6. for explanation and simulation/experimental results.

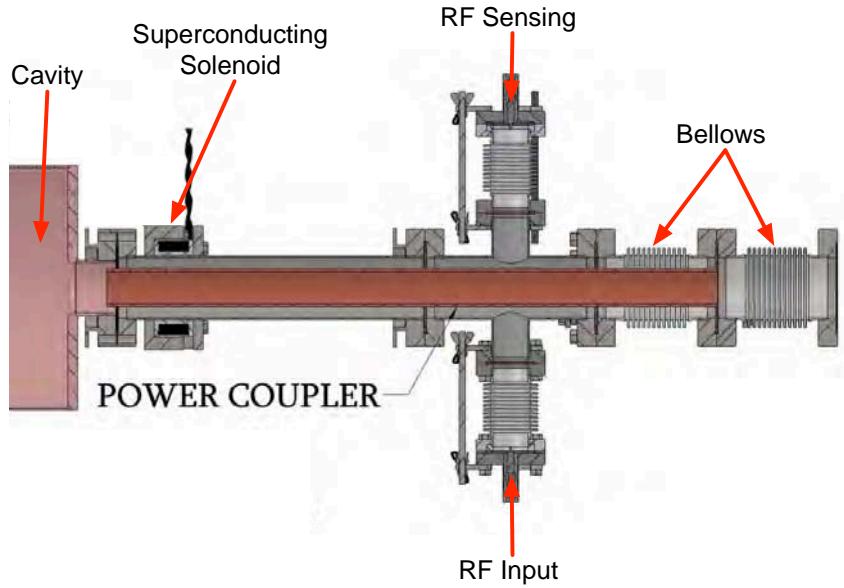


Figure 20. The NPS cavity RF coupler drawing. The pink portion on the left is the 500 MHz SRF cavity. The superconducting solenoid is shown near the coupler tip. The vertical structures are input and output antennae for exciting the waveguide and for reading the return RF signal from the cavity.

antenna. This is done through capacitive coupling. The connection directly opposite of the RF input connection is a receiving antenna, also capacitively coupled. This antenna receives two signals: one is a direct path transmitted signal from the input antenna and the other is a signal from the cavity that is proportional to the power contained in the cavity. As the cavity fields “fill,” this return signal is small. Once the cavity has reached its maximum capacity, the return signal reaches a steady state level.

This coupler configuration also allows for adjustment of the coupling constant to the SRF cavity. By compressing/extending the bellows around the coupler attachment to the beam tube wall, the coupler tip location relative to the cavity can be adjusted. The coupling constant is a measure of the overlap between the evanescent coupler fields and the cavity fields [12]. Ideally, the coupler is adjusted such that “unity coupling” is achieved. Unity coupling indicates that the cavity is acting as a

matched load to the coupler, allowing for efficient transfer of power from the coupler to the cavity. To measure the coupling constant, one can observe the behavior of the transmitted and return power levels of a square wave gated RF pulse on a frequency analyzer. Figure 21 shows illustrative plots for this concept.

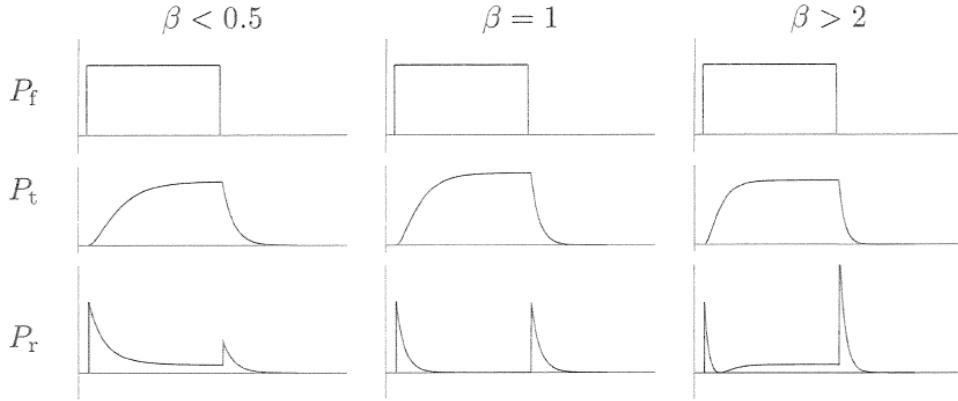


Figure 21. The left column shows weak coupling. P_f is the RF feed power, P_t is a measure of the power contained in the cavity, and P_r is the power reflected back to the feed, but also contains the power emitted from the cavity. When $\beta = 1$, we note that the emitted power and reflected power exactly cancel, while for $\beta > 1$, the emitted power dominates and for $\beta < 1$ the reflected power dominates P_r . The first peak in P_r is associated with the ratio of reflected to forward power and the second peak is associated with the ratio of emitted to forward power. After [12].

The coaxial coupler antenna is detuned from the fundamental frequency of the cavity, 500 MHz, so that it acts as a transmission line rather than a resonator. The coupler has a length of 60.5 cm corresponding to resonant frequencies of 123.9 MHz ($\lambda/4$), 371.6 MHz ($3\lambda/4$), and 619.4 MHz ($5\lambda/4$). Driving the antenna at 500 MHz allows for RF coupling to the cavity without the antenna being resonant. The input and sensing antennae are located away from the shorted end of the transmission line in order to enable driving the RF amplitude on the coupler as well as allowing for reception of the return signal on the output antenna. However, with the required drift length of about 75 cm from the cavity (this length includes the bellows section after

the coupler termination), there is significant opportunity for electron beam growth and emittance degradation at low energies. While the superconducting solenoid provides focusing, from a beam line layout perspective, it would be advantageous to position a booster cavity as close as possible to the gun. In the booster configuration, the level of desirability of having another booster cavity located close to this one is dependent on the beam energy prior to entering the 500 MHz booster.

In actual practice, this coupler arrangement turns out to be extremely sensitive and difficult to diagnose. In the simpler booster configuration the coupler arrangement is really three coupled systems: the input antenna, the coaxial transmission line, and the SRF cavity. An adjustment to one of the antennas does not have easily predictable results on the entire system and the arrangement is not easily modeled for predictive simulation purposes. In the experiments reported later, the input antenna was inserted in as far as possible without shorting to the coaxial line and all antenna adjustments (to adjust β) were made by changing the compression of the coaxial coupler termination bellows.

In the next iteration of this cavity, it is highly recommended that an intracavity loop antenna or transverse coaxial coupler be used, an example of which is shown in Figure 22. A rectangular waveguide at the drive frequency desired would be quite large relative to the small diameter pipe connecting to the cavity, making it less practical. An intracavity loop antenna requires an access port through the cavity wall requiring a vacuum tight weld. While introducing a penetration complicates the construction of the cavity and cryomodule and introduces a possible increase in the number of surface defects and their associated effects on cavity performance, the increase in control of the RF coupling and decrease in beamline length may compensate these drawbacks. The transverse coaxial coupler allows for greater RF power input than the antenna version while maintaining the cavity with only beam tube penetrations and achieving a much reduced beamline length over the current design.

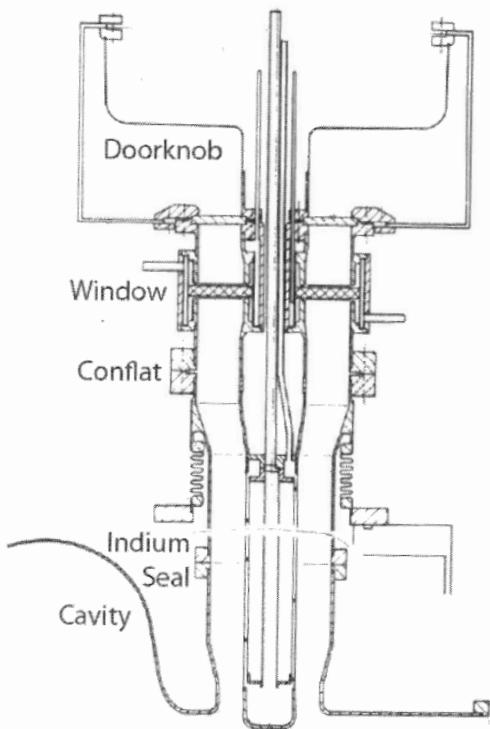


Figure 22. A transverse coaxial coupler utilizing a similar coupling mechanism (although waveguide fed in this case) that could be used to provide power to the NPS cavity while greatly reducing the drift length to the next beamline element. The antenna is the central vertical structure that projects slightly into the beam tube. From [12].

3. Cathode

To take advantage of the benefits of the cavity design, the cathode assembly required a significant amount of design work. As the gun is intended as an experimental set-up for multiple cathode types and materials, the assembly must be removable to exchange cathodes for testing. In most guns, this process is handled by a load-lock system allowing for insertion/removal of a cathode without breaking vacuum in the gun or cathode assembly, [24] for example. As these load-lock systems are very expensive and one-off designs, and the cavity cannot be tested in the gun configuration to verify simulation until some form of cathode is installed, it was determined that using a standard metal cathode was a suitable, low risk first step while the remainder of the cathode assembly interface requirements were fleshed out with other research groups. To prevent foreign metal contamination in the prototype cavity, we chose to use niobium as the cathode material. It has a work function similar to copper ($\text{Nb} \simeq 4.3 \text{ eV}$, $\text{Cu} \simeq 4.7 \text{ eV}$ [6]). In order to liberate electrons via the photoelectric effect, the incoming photon must have energy equivalent to the materials work function; for niobium, this requires a laser with a wavelength shorter than 288 nm. Niobium and copper also have similar quantum efficiencies, on the order of 10^{-5} , depending on cathode preparation techniques used. As we would also like to test some of these cathode preparation techniques[25], [26], [27] and there exists the possibility for ablation of cathode material during cathode processing, the niobium cathode material choice presents the least damage risk to the cavity. Using another metal could result in plating of that material inside the cavity, thus, degrading the superconducting properties of the cavity. The use of an alkali metal on a semiconductor substrate requires better vacuum levels than currently available on the system. Alkali metal cathodes also present issues (possible contamination of the cavity) that we wish to avoid until the cavity has been properly characterized.

Once it was decided that the cathode would be metal, the next major decision was how the cathode should interact with the cavity electrically. One possibility

is to mechanically short the cathode stalk to the cavity, preventing the RF power from passing down the cathode stalk out of the cavity. Recognizing that one of the inherent benefits of the cavity's geometry is the adjustable radial fields on the cathode surface, any simple RF short would inhibit the use of this characteristic, making this approach less desirable. Additionally, any short would require physical contact in the high field region of the cavity. Any metal particulate introduced here has a very high probability of generating a field emission site, degrading gun performance. The other possibility is a cantilevered stalk, not shorted directly to the cavity, to support the cathode in the desired position. The simplest non-shorted cathode stalk is a right cylinder inserted into the nose cone area.

This simple design comprised the first iteration of the cathode stalk. It became apparent quickly that there was a singular flaw with this design. Just like the input coupler, the cathode stalk becomes an RF transmission line allowing RF energy to flow down it as a coaxial waveguide. A simple calculation can show how much energy this configuration will pull from the cavity and turn into resistive heating. This heat is a significant concern because the cathode stalk is not ideally connected to the outside environment for conductive cooling, and without active cooling, the cathode stalk could present a non-trivial heat load to the cryogenic system, possibly impacting the ability of the cavity to remain superconducting in the gun configuration. Additionally, any RF power pulled from the cavity decreases the cavity fields used for beam acceleration unless additional RF power is supplied (which would also increase the power lost to the cathode stalk).

To model this system analytically, we simplify the cavity to a pillbox with a cathode set at the input side of the cavity. The cathode stalk extends inside a vacuum pipe to a short at the base where it connects to the beam pipe, as seen in Figure 23.

The cavity fields cause a surface charge, sourced from the stalk, to build up on the cathode. We can determine this charge using Gauss' law, forming a Gaussian pillbox around the cathode surface. If we place the pillbox boundary near the surface,

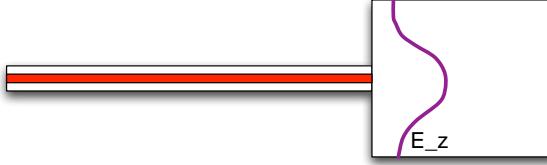


Figure 23. A resonant cavity has E_z as shown. This field impinges on a cathode attached to a right cylindrical stalk. The cavity resonates at a fundamental frequency, f_0 .

the non- \hat{z} components can be neglected, leaving only $E_0\hat{z}$. We calculate the surface charge

$$\oint_{\mathcal{S}} \mathbf{E} \cdot \hat{\mathbf{n}} \, dA = \frac{q}{\epsilon_0},$$

where \mathcal{S} is the surface of the Gaussian pillbox, A is the area of the cathode surface, $\hat{\mathbf{n}}$ is the outward normal vector of the Gaussian pillbox, and q is the charge enclosed by the surface. Solving this equation, we find the charge that builds up on the cathode surface is

$$q(t) = E_0 \epsilon_0 A \sin(\omega t) , \quad (\text{III.14})$$

where we have included the oscillatory nature of the electric field with amplitude E_0 and angular frequency $\omega = 2\pi f_0$. If we take the derivative of (III.14), we get a current

$$\frac{dq}{dt} = E_0 \epsilon_0 A \omega \cos(\omega t) = I_0 \cos(\omega t) , \quad (\text{III.15})$$

where $I_0 \equiv E_0 \epsilon_0 A \omega$. This current is conductively attached to the cathode assembly, which acts as a coaxial transmission line relative to the cavity. The current that populates the cathode “sloshes” along exterior surface of the stalk, and we can state that the current density at each point along the stalk (assuming the stalk is approximately $3\lambda/4$ relative to the cavity fundamental so that the stalk couples well to the cavity) is

$$J_{stalk}(z, t) = \frac{I_0}{2\pi R} \cos(\omega t) \cos\left(\frac{2\pi z}{3\lambda}\right) . \quad (\text{III.16})$$

We integrate out the z dependence in preparation for finding the power,

$$\begin{aligned}
I_{stalk}(t) &= \frac{I_0}{2\pi R} \cos(\omega t) \int_0^{\frac{3\lambda}{4}} \sin\left(\frac{2\pi z}{\lambda}\right) dz \\
&= \frac{I_0 \cos(\omega t)}{2\pi R} \left[\frac{-\lambda}{2\pi} \cos\left(\frac{2\pi z}{\lambda}\right) \right]_0^{\frac{3\lambda}{4}} \\
&= \frac{I_0 \cos(\omega t)}{2\pi R} \left(\frac{\lambda}{2\pi} \right) \\
&= \frac{I_0 \lambda \cos(\omega t)}{4\pi^2 R}.
\end{aligned} \tag{III.17}$$

Since we desire to find the average power lost to the cathode stalk and we have the current already, we use the ohmic power equation

$$\langle P \rangle = \langle I_{stalk}^2 \rangle \mathcal{R}. \tag{III.18}$$

Starting with the resistance, we know

$$\mathcal{R} = \frac{\rho l}{A_\delta}, \tag{III.19}$$

where ρ is the resistivity of the material, l is the length of the conductive path, and A_δ is the this annulus shaped cross-sectional area through which current is flowing. The cross-sectional area in this case is a small ring at the outer surface of the stalk whose depth is equal to the skin depth of the material, δ , at the driving frequency, f_0 . Thus the area is

$$A_\delta = \pi R^2 - \pi (R - \delta)^2, \tag{III.20}$$

where R is the radius of the cathode stalk and

$$\delta [\text{m}] \approx 503.3 \sqrt{\frac{\rho}{f_0}},$$

where f_0 has units of [Hz] and ρ has units of [$\Omega \text{ m}$]. Substituting (III.20) into (III.19), we obtain

$$\mathcal{R} = \frac{\rho \lambda}{\pi \delta (2R - \delta)}, \tag{III.21}$$

where l is assumed to be a longer length, $\lambda = c/f_0$, for power dissipation purposes.

Returning to equation (III.18), we now have

$$\begin{aligned}
\langle P \rangle &= \left\langle \frac{I_0 \lambda \cos(\omega t)}{4\pi^2 R} \right\rangle \frac{\rho \lambda}{\pi \delta (2R - \delta)} \\
&= \frac{I_0^2 \lambda^2}{32\pi^4 R^2} \frac{\rho \lambda}{\pi \delta (2R - \delta)} \\
&= \frac{E_0^2 \epsilon_0^2 A^2 \omega^2 \lambda^3 \rho}{32\pi^4 R^2 [\pi \delta (2R - \delta)]} \\
&= \frac{c^3 E_0^2 \epsilon_0^2 R^2 \rho}{8\pi f_0 \delta (2R - \delta)}.
\end{aligned} \tag{III.22}$$

Choosing some representative numbers for the gun design: $E_z = 20$ MV/m, $f_0 = 500$ MHz, $R = 1$ cm, and room temperature copper as the stalk material, we find that the power dissipated in the stalk is 1.94 kW. Comparing this with a Superfish model of the actual NPS cavity with comparable parameters and a right cylindrical cathode stalk, we find 2.04 kW dissipated in the cathode stalk.

Analyzing equation (III.22), we see that there are several parameters that can be used to control the power loss along the stalk. Reducing fields on the cathode surface, decreasing the operating frequency (which also increases the skin depth and further decreases the power), choosing a material with a lower resistivity, and decreasing the area of the cathode surface area (controlled by the stalk radius). Of these, the frequency of the cavity is fixed; however, the material choice for the cathode stalk, cathode surface area, and cathode surface fields are all adjustable within the confines of the cavity design.

Additional losses can be attributed to RF power being coupled onto the cathode stalk as a transmission line. This problem is not a new one and has been observed/predicted in both the Rossendorf gun and the proposed Brookhaven SRF gun. Both of these guns were envisioned as guns from the outset and were designed to incorporate an RF choke, essentially a resonant cavity that preferentially gathers the RF power and prevents it from flowing down the cathode stalk. Figure 24 shows the choke designs incorporated into these guns.

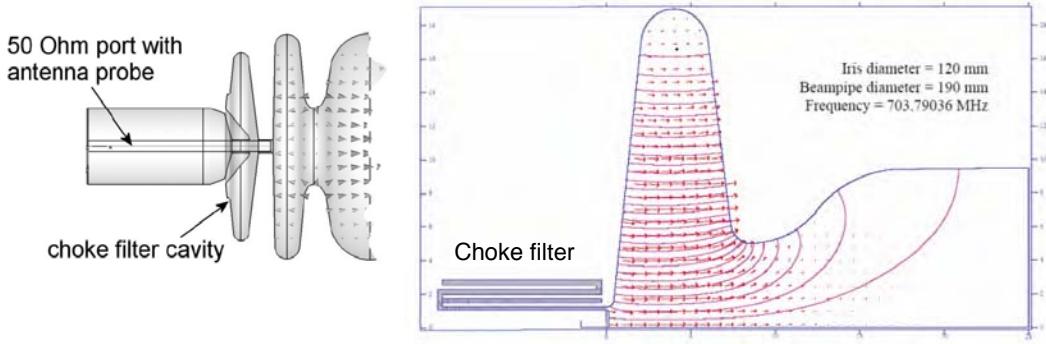


Figure 24. RF choke designs for two SRF gun designs. On the left is the Rossendorf gun and the choke cavity just behind the cathode surface. From [15]. On the right is the proposed BNL/AES SRF gun with a different choke design placed very near the cathode surface. After [11].

Since the NPS design was initially contemplated as a booster, there is no RF choke incorporated in the superconducting section of the cavity. To reduce RF power flowing down the cathode stalk, our approach, proposed by our Niowave collaborators, has been to shape the cathode stalk so that the cathode stalk acts as a Bragg reflector, greatly reducing the RF power coupling down the cathode stalk. Figure 25 shows the final version of the cathode assembly. Note that it incorporates several features to reduce RF coupling. The cathode holder has a narrow fit into the nose cone to reflect as much RF power back into the cavity as possible. The multiple diameter changes along the support stalk act as impedance mismatches (or a Bragg reflector) to reflect any power that has passed down the stalk back toward the cavity. In addition to changing the stalk geometry from the simplified model, we explored the impact of changing the conducting material inside the beam tube and the application of active cooling. These strategies were modeled in Superfish and the results are summarized in Table 3.

The geometry changes alone decrease the losses sufficiently to allow for operation as a gun. Should further heat reduction become necessary, additional measures can be taken. Anodizing the interior of the vacuum pipe with a layer of copper and no

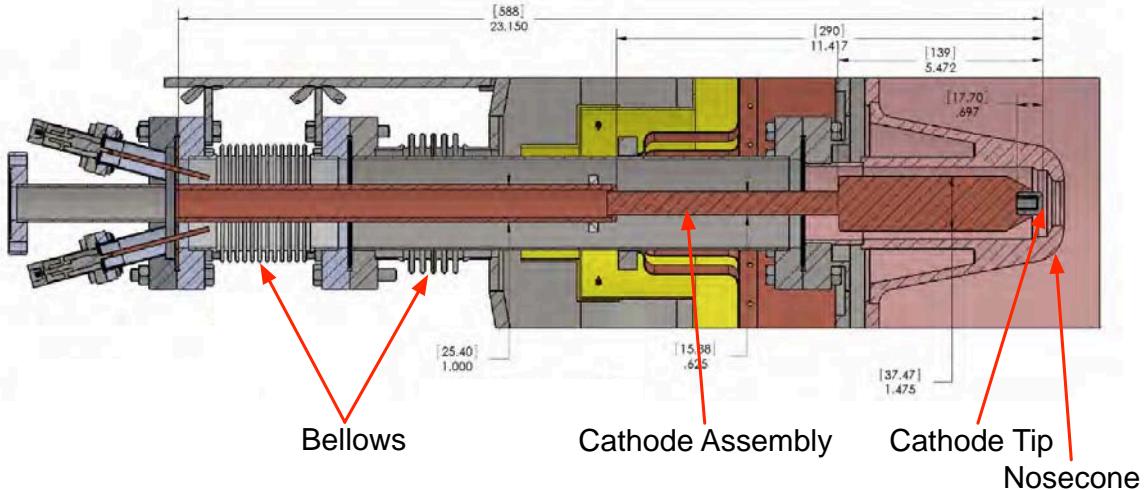


Figure 25. The assembly used for insertion and positioning of the cathode surface in the NPS SRF cavity gun configuration.

further changes results in a reduction in ohmic heating to half of the original design. The thickness of this copper coating would need only be $10 \mu\text{m}$ to achieve a thickness greater than 3 skin depths at 500 MHz. If active cooling of the cathode assembly is desired, making the tip end hollow and passing a liquid nitrogen feed would reduce the radiated power to the cavity significantly as well as decreasing the material's resistivity. Active cooling could also be applied to the exterior of the vacuum pipe as well, although the effect would be smaller than cooling the cathode stalk. Applying all the improvement measures considered results in a reduction of ohmic losses to 2.7% of the original design.

One of the best attributes of this cathode/cavity combination is the ability to tune the fields on the cathode surface by adjusting the cathode's position relative to the nose cone. The bellows assembly at the shorted end of the cathode assembly allows for the longitudinal positioning of the cathode surface from a position 1.51 mm beyond the nose cone (with fields $E_{z, \max} = 38 \text{ MV/m}$ and $E_{r, \max} = 6.8 \text{ MV/m}$ defocusing at $r = 5 \text{ mm}$ measured 1 mm in front of the cathode face) to 14.23 mm retracted behind the nose cone (with field $E_{z, \max} = 5.8 \text{ MV/m}$ and $E_{r, \max} = 483$

Table 3. Cathode assembly and beam pipe power dissipation due to coupling with the NPS SRF cavity as calculated by Superfish. Various combinations of cathode stalk design, materials (304 stainless steel and copper), and active cooling are considered. The simple model uses a right cylindrical cathode stalk. The complex model uses the changing radius cathode stalk. For each model, power dissipated in the vacuum tube wall (pipe) and along the cathode stalk (K) are shown.

Component Arrangement	Simple Model		Complex Model	
	Pipe [W]	K [W]	Pipe [W]	K [W]
SS Pipe (293 K), Cu K (293 K)	3061	2043	121.5	334.1
SS Pipe (77 K), Cu K (293 K)	1190	2043	47.21	334.1
SS Pipe (77 K), Cu K (77 K)	1190	794.1	47.21	129.9
Cu Pipe (293 K), Cu K (293 K)	469.0	2043	18.61	334.1
Cu Pipe (77 K), Cu K (293 K)	182.3	2043	7.234	334.1
Cu Pipe (77 K), Cu K (77 K)	182.3	794.1	7.234	129.9

kV/m (defocusing) at $r = 5$ mm and -58 kV/m (focusing) at $r = 2.5$ mm measured 1 mm in front of the cathode face). Figure 26 shows these two cathode position extrema and the associated field lines as calculated by Superfish. It can easily be seen between Figure 26 and Figures 17 and 18; there is a distinct trade between the main accelerating gradient and the focusing fields experienced by the beam as it is emitted from the cathode surface. Figure 26 also plainly shows the cathode face is a blunt, flat-faced surface. Additional focusing during electron emission could be obtained by dishing this surface.

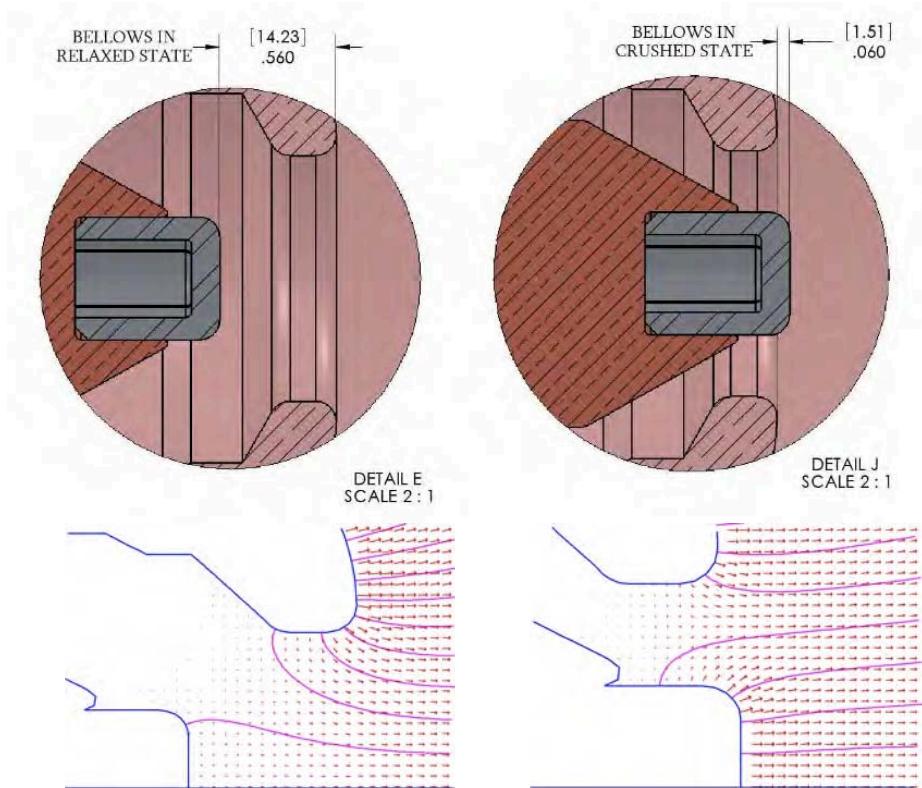


Figure 26. By compressing or extending the cathode assembly bellows, the cathode position relative to the nose cone can be adjusted. This changes the radial and longitudinal fields experienced across the cathode face. The upper diagrams are from Niowave design drawings. The bottom diagrams are field plots of the nose cone/cathode region of the cavity taken from Superfish.

4. Superconducting Solenoid

A superconducting solenoid is included in the cavity design as a focusing element, mounted as close as practicable to the SRF cavity. Figure 20 shows the relative position of the solenoid to the cavity. As the electron beam exits the cavity, space charge effects will be causing the beam to expand radially. By placing the solenoid as close as practicable to the SRF cavity, we maximize the bunch charge that can be extracted through the cavity coupler.

The solenoid is chosen to be superconducting as it is located near the helium dewar and inside the liquid nitrogen shield. A heat source at this location would be a significant heat load on the cryogenic cooling systems. Using a niobium-titanium winding (niobium wires in a titanium matrix) and conductive cooling from the helium dewar, when the wire temperature is below the critical temperature for niobium, a DC current can be passed through the solenoid with negligible ohmic heating. Conductive cooling is provided to the solenoid casing through two copper bus bars, one of which is shown in Figure 27.

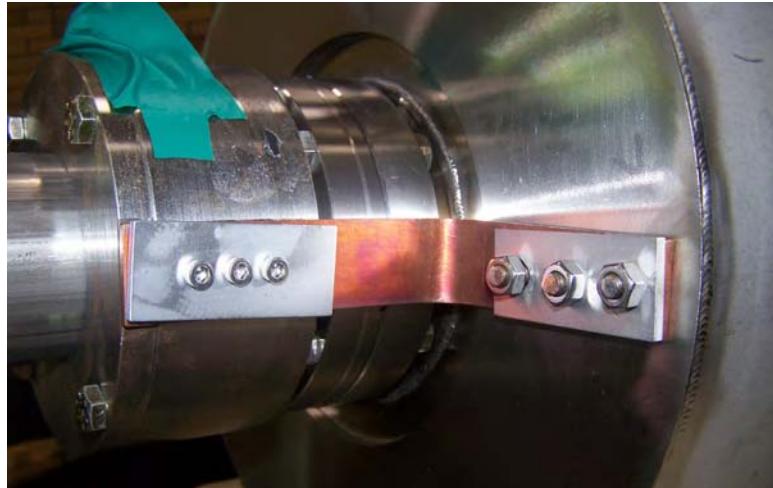


Figure 27. Copper bus bar providing conductive cooling from the liquid helium dewar to the superconducting solenoid. There are two identical bars with the second one on the opposite side of the beam pipe from the one shown.

To prevent magnetic fields from becoming trapped in the accelerating cavity, the solenoid's iron flux clamp will be demagnetized after use and the solenoid will only be energized after the cavity has reached superconducting temperatures. The solenoid is designed for a peak field of 238 mT with a current of 5.0 A. The solenoid, after manufacture, was tested to handle a maximum current of twice the design value without quench.

5. Assembly and Cavity Cleaning

Once the cavity pieces have been manufactured and machining completed, they must be assembled in a fashion that minimizes possible defects affecting cavity performance. Once the cavity is assembled, the difficulty in correcting any defects that arise increases drastically. Defects that can affect cavity performance include weld spatter, dust, metal shavings, and hydrogen contamination, among many others.

a. *Cavity Assembly*

The assembly process starts with welding the cavity into its final state. All preliminary testing to verify the machining steps⁴ must be completed prior to this step, as disassembly of the cavity is extremely difficult and costly. Normal welding processes can be used, but experience with processes such as tungsten inert gas (TIG) welding has been poor [12]. Electron beam welding offers a much more controllable environment for ensuring weld quality and minimizing incorporation of gases in the weld bead. Figure 28 shows the NPS cavity in an electron beam welder at Sciaky, Inc., in Chicago, IL.

The weld procedure must be well planned to ensure all welds are full penetration and have a smooth surface at the cavity inner surface. The process must be controlled to prevent detuning of the cavity due to weld shrinkage (excessive heat)

⁴See Chapters IV.B.1. and IV.B.2. for results of these steps.



Figure 28. The NPS cavity in the electron beam welder at Sciaky, Inc. Top left is the nose cone being welded to the base plate. Bottom left is the base plate being welded to the outer conductor of the cavity. The photo on the right is a conflat flange being welded to the anode side of the cavity.

or blow out (focused electron beam). Electron beam welding is conducted under vacuum usually less than $10 \mu\text{Torr}$ [12], which minimizes the gases present to absorb into the cavity materials and welds.

b. Cavity Surface Etching

The manufacturing process introduces contaminants in the form of oxidation layers, foreign metal contamination, and lubricants that must be removed for optimal cavity performance. Removal of lubricants is best accomplished through use of the usual solvents used in vacuum systems. Accomplishing the removal of the oxidation layer and foreign metals require more aggressive methods. The standard method adopted for this process is the buffered chemical polish (BCP). BCP is an acid mixture consisting of a solution of hydrofluoric acid (HF), nitric acid (HNO_3),

and phosphoric acid (H_3PO_4) in a ratio of 1:1:2 respectively. Figure 29 shows the setup used to perform the BCP on the NPS cavity.



Figure 29. This station was used for performing the buffered chemical polish of the NPS cavity. In the inset at lower right is the cavity placed in the rocking cradle that mates with the stand shown just to the right of the cooling tub. The hole shown in the inset is the connection to the liquid helium fill tube and leads to the exterior of the niobium cavity.

All acids were premixed and maintained under the vent hood during the processing. The acid was connected to the cavity via plastic tubing attached to the beamline connection points (white connectors in Figure 29 inset) including a drain line. The cavity was filled in the vertical position with the nose cone end down and drain line open at the top to ensure proper filling of the entire cavity. Once filled, the drain and fill lines were clamped and the cavity placed in the rocker

frame which was in the cooling tub. The cooling tub was filled with a mixture of water and ice to provide cooling of the cavity during the etching process. A thermometer was used to measure the temperature of the acid solution inside the cavity and maintain it below 18 °C to control the etch rate and minimize the chance of hydrogen contamination [12]. During the etch, the cavity was agitated by hand to prevent settling of niobium removed from the cavity surfaces. The etch was halted after the allotted time sufficient to etch approximately 150 μm of surface material by dumping the acid followed immediately by flushing with ultrapure water (resistivity of 17.5 M Ω -cm as measured at the discharge of the storage tank). After flushing, the cavity was filled with ultrapure water and capped for transport to the clean room to prevent any airborne particulates from having access to the cavity.

c. High-Pressure Rinse

Once inside the class-100 clean room, the cavity was drained and mounted to the high-pressure rinse station shown in Figure 30. As has been seen in previous cavities, high-pressure water rinse immediately following cavity etching or electropolishing has shown significant reduction in field emission allowing for attainment of higher field gradients [28]. The water rinse removes any last traces of the acids used to chemically process the cavity as well as flushing any remaining particulate matter from the cavity walls. This must be performed in a “clean” environment to prevent any particulates in the air from adhering to the liquid surface and being deposited as the liquid evaporates, negating the benefits of the high-pressure water rinse.

The wand consists of a single, small pipe with a series of small holes drilled near the tip, perpendicular to the surface. The tip of the wand is angled toward the exterior of the cavity and has a nozzle drilled in it to provide an angled spray as well. High-pressure water is provided at the opposite end of the wand connected to a high-pressure pump outside the clean room and taking suction from the ultrapure water tank. During rinsing, the wand is cycled up and down while the cavity is simultaneously rotated. The cavity is mounted with the nose cone at



Figure 30. The Niowave high-pressure rinse station consists of a high-pressure, ultrapure water feed to a rinse wand. One the left is the stand in the vertical position without the cavity showing the wand in the fully extended position. On the right is the cavity mounted to the station with the wand in the angled position to allow for cleaning of the exterior of the nose cone surface.

the top (pointed toward the mounting ring). This allows for canting of the wand to reach into the rear corners of the cavity and clean the exterior of the nose cone and coaxial line. These positions are of utmost importance as they see the highest electric and magnetic fields. The cavity was rinsed in each position continuously for approximately 1.5 hours and then allowed to air dry in the clean room. Once the majority of the water had evaporated, the cavity was placed under vacuum to verify that the cavity was still vacuum tight and to prevent any particulate intrusion in the now clean and prepared cavity. The cavity at this stage is shown in Figure 31 with the superconducting solenoid attached to the beam tube and wired for final installation.

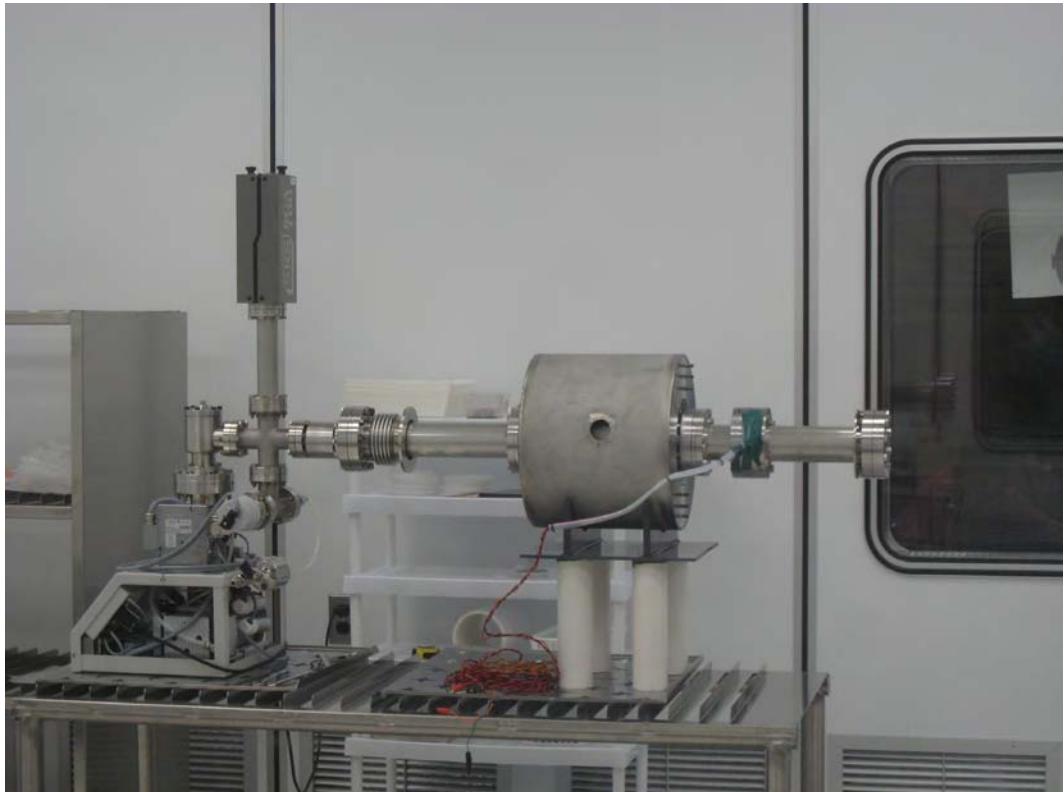


Figure 31. The NPS SRF cavity in the class-100 clean room after chemical processing and high-pressure rinse. The cavity is under vacuum to remove the remaining water and to verify the cavity is still vacuum tight. Note that the beam tubes have also been attached to the cavity in preparation for the next stage of assembly.

d. Cryomodule Assembly

The cavity and beamline pipes were then brought back to atmospheric pressure under dry nitrogen purge and the beam tubes capped with blank flanges so that the cavity could be encapsulated in the cryomodule. The cryomodule consists of multiple layers of insulation and cryogenic shielding in order to minimize heat transport from the outside environment to the liquid helium. As part of the assembly process, multiple temperature sensors were installed to monitor temperatures in the various parts of the cryomodule. The most important sensors are those located on the solenoid casing and at the top of the dewar. The first indicates whether the solenoid has reached superconducting temperatures and the second can be used as a liquid helium level sensor. Each sensor is wired to a remote monitoring device for presentation to the operators. The wiring and temperature monitors are shown in Figure 32.

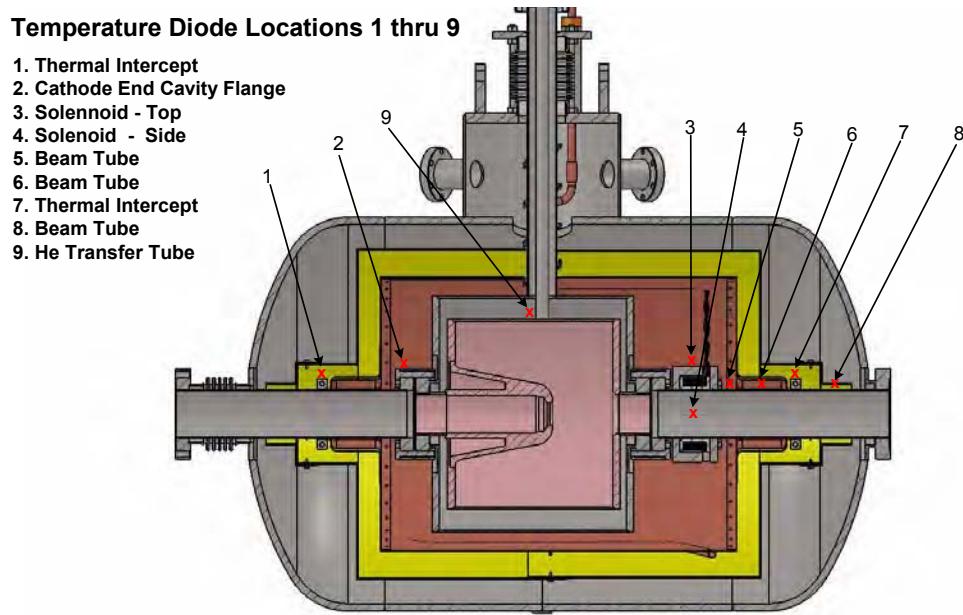


Figure 32. Nine temperature monitoring sensors are placed inside the cryomodule to monitor temperatures of the cavity and support structures. These sensors read out to a remote monitoring station for use by operations personnel.

Once prepared, the cavity was moved into place under a support structure for the “conning tower.” The conning tower supports the cavity inside the cryomodule via the helium fill tube, allowing the cavity to hang from the conning tower. The conning tower is welded to the main portion of the cryomodule to complete the vacuum tank, providing sufficient support for the cavity. The cavity is hung from a truss for alignment relative to the support foundation for assembly of the remaining portions of the cryomodule as seen in Figure 33.



Figure 33. The cavity is in place and aligned to the support foundation.

After a test fitting of the various portions of the cryostat, the cavity was enveloped step by step. The first thermal layer after the helium dewar is three thicknesses of insulating blanket. The insulating blanket material is shown in Figure

34 and consists of a layer of expanded fabric backer attached to a metalized mylar surface. The backer's expanded structure limits conductive thermal transport while the metalized surfaces act to reflect radiated heat from the outer layers of the cryostat.



Figure 34. Insulating blanket material (shown on left) is wrapped around the helium vessel.

Surrounding the insulating blanket, a liquid nitrogen shield is installed. The liquid nitrogen shield, shown in Figure 35, is made from copper with copper tubing brazed to the exterior for conductive cooling of the shield to liquid nitrogen temperatures. During operation, liquid nitrogen is constantly flowed through the shield to maintain it at 77 K. To ensure the liquid nitrogen system is kept liquid throughout, a constant drip is maintained from the outlet where it is observable by the operators. Reducing the temperature decreases the radiated power incident on the insulating blankets by a factor of >230 over room temperature.

To minimize stray magnetic fields affecting the cavity's superconducting properties, the entire helium vessel is encased in a MuMetal shield: see Figure 36. MuMetal is a nickel-iron alloy ($\approx 80\% - 20\%$) with small amounts of molybdenum that has an extremely high magnetic permeability [29]. The MuMetal acts to shield the SRF cavity from the earth's magnetic fields and any stray magnetism in the steel vacuum vessel or local stray fields from the SRF cavity. This is extremely important, as magnetic field lines can become "pinned" within the niobium as the SRF cavity



Figure 35. The copper liquid nitrogen shield in place around the cavity on the left. On the right is shown the attachment of positioning strips to keep the cavity suspended within the shield and minimizing conductive thermal contact. The black covering seen on the left was in place only for initial fit testing and was removed for final installation (removed on right).

goes through transition. These trapped flux lines then contribute to the cavity's residual resistance, increasing the ohmic losses and limiting cavity performance.

Surrounding the MuMetal and liquid nitrogen shield is a nonmagnetic vacuum vessel. This vessel, shown in Figure 37, provides structural support for the cavity and beam tubes as well as providing significant thermal isolation from the lab environment. The vacuum vessel is welded to the collars at each end of the beam tubes connecting to the cavity. In case rework or reprocessing of the cavity is necessary, this encapsulation can be removed by removing a welded band and cutting the skip welds holding the two halves together. While simple sounding, the effort required to remove the cavity from its shields and rebuild the cryomodule is significant and should be avoided if at all possible.

Completing the vacuum vessel is the “conning tower” shown prior to assembly in Figure 38. The conning tower is the access point for the cryogen fill and exhaust as well as the temperature sensors. Additionally, the conning tower houses



Figure 36. MuMetal shield installed around the liquid nitrogen shield. Note that superconducting solenoid is encapsulated by MuMetal. The black covering indicates this picture was taken during test fitting.

ports for the guard vacuum burst disk, vacuum pumping, and vacuum monitoring. The insulating vacuum for this cryomodule is separate from the beamline vacuum. Vacuum levels in the vessel are maintained at 10^{-4} to 10^{-5} Torr.

To support the cavity when warm and during eventual shipment from Niowave to NPS, two small stands are installed. These stands allow the helium dewar to rest upon them when warm, preventing the cavity from swinging on the helium fill tube and exit beam pipe. The feet are small hollow tubes to minimize heat conduction through the cryomodule's thermal shields (see Figure 39). When the cavity is cooled to liquid helium temperatures, the contraction of the fill tube is sufficient to lift the helium dewar from the feet which, in addition to removing the thermal path to the lab environment, also provides some small isolation of the cavity from external sources of mechanical vibration.

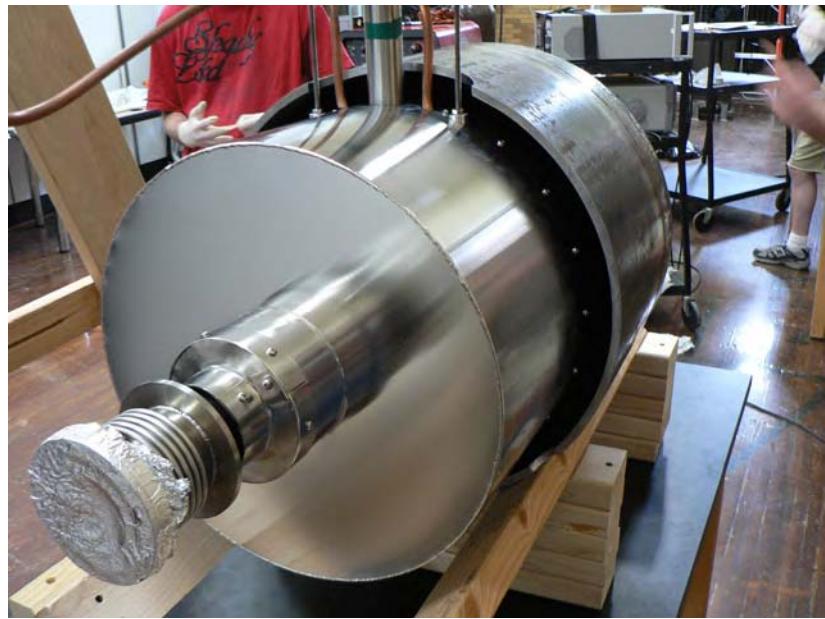


Figure 37. The vacuum vessel is arranged around the cavity and shields.



Figure 38. The conning tower assembly prior to assembly and mating to the vacuum vessel.



Figure 39. When the cavity is warm, it rests upon two “shipping feet” that support the cavity in addition to the liquid helium fill tube and exit beam tube. When cold, the contraction of these tubes lifts the cavity from these supports, isolating it thermally and mechanically.

e. Coupler Installation

Once the cryomodule construction was complete, the entire cryomodule and foundation returned to the clean room for installation of the RF coupler. The RF coupler was cleaned and assembled separately. The coupler was kept under vacuum in a surrogate beam tube for cleanliness purposes until the cryomodule was ready to receive the coupler. Figure 40 shows the coupler and pumping station ready to be mated to the cryomodule. The pumping station includes a leak-off valve for introducing gases into the vacuum, a residual gas analyzer, and an “up to air” valve for connection to a dry nitrogen purge when vacuum must be released. The vacuum pump is a combined roughing and turbo pump capable of maintaining vacuum in small volumes of 10^{-9} torr. What is seen in Figure 40, in combination with the cryomodule, constitutes the entire system for the initial RF testing.



Figure 40. The coupler was maintained in a clean environment, under vacuum, while the cryomodule was assembled. The coupler is shown on the left. The vacuum pumping stand is on the right.

After re-entry into the clean room, the cryomodule and foundation were thoroughly cleaned and prepared for mating with the coupler assembly. Using standard clean room procedures to minimize particulate generation, the surrogate

beam tube was removed from the coupler and the output beamline of the cryomodule was opened. Figure 41 shows the coupler after the surrogate beam tube is removed. The coupler is polished, annealed copper cleaned with ultrapure water prior to being placed under vacuum.



Figure 41. The RF coupler assembly with surrogate beam tube removed. The tube is supported at the visible flange by a teflon spider.

The coupler stand is set upon teflon rails of the same thickness as the aluminum base plates that will mount the coupler stand to the foundation. Teflon rails are used to facilitate manual movement of the coupler into the cyromodule beam tube while minimizing particulate generation. Any particulate contamination of the cavity will increase the risk of degraded performance due to field emission sites. Particulate contamination of the coupler could result in multipacting, limiting the obtainable cavity fields. Figure 42 shows the insertion process. Technicians must ensure the coupler remains centered in the beam pipe and parallel to the axis during insertion to prevent impacting the beam tube and possibly damaging the coupler surface. Once in place, the flange between the sections is tightened, and the coupler assembly and

cavity are placed under vacuum. Final alignment of the coupler requires RF power and will take place outside of the clean room. Until the cathode assembly is ready for installation, the upstream end of the cryomodule will remain closed with a blank flange.



Figure 42. Insertion of the RF coupler into the cryomodule beam tube.

After leaving the clean room, the cryomodule received a coat of paint and remaining ancillary systems were connected to prepare for initial RF testing.⁵ Figure 43 shows the Niowave facility set up for initial RF testing of the cavity. Figure 44 shows a close up of the cryomodule and beamline with the ancillary equipment connected. On the left is the pumping station as described previously. The metal shroud around the shorted end of the coupler is supporting structure for compressing and extending the bellows section for adjustment of the coupling strength to the cavity. Under the RF coupler is the input RF power feed from a 100 W amplifier. The amplifier is driven by a 500 MHz signal generator controlled at the operator station.

⁵Initial RF testing results are discussed in Chapter IV.B.4

Directly under the cryomodule is an ionization chamber for measuring radiation generated during RF testing. The black tubes coming from the top of the cryomodule are insulated liquid nitrogen lines connecting to a liquid nitrogen tank seen in Figure 43.

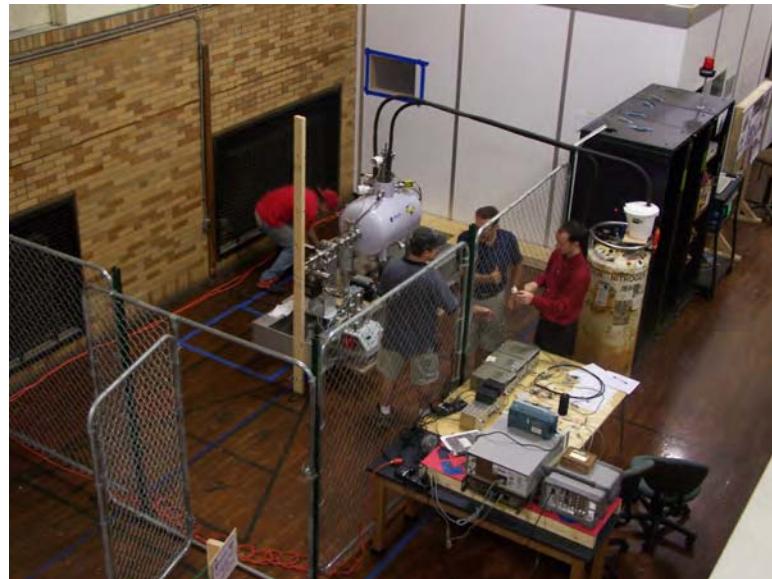


Figure 43. Niowave facility ready for RF testing of the NPS 500 MHz gun/booster.



Figure 44. The completed 500 MHz SRF gun/booster ready for RF testing.

f. Cathode Assembly Installation

After initial cryogenic and RF testing, the cathode assembly was installed in preparation for further RF processing and first beam test. Prior to installation, the cathode assembly was cleaned, and the cathode stalk placed in a surrogate beam pipe while in the clean room to maintain cathode cleanliness. The cathode end of the cryomodule was encapsulated in a portable clean room to minimize introduction of particulates into the cavity during the cathode assembly installation process. Figure 45 shows the cathode assembly inside the portable clean room prior to installation.

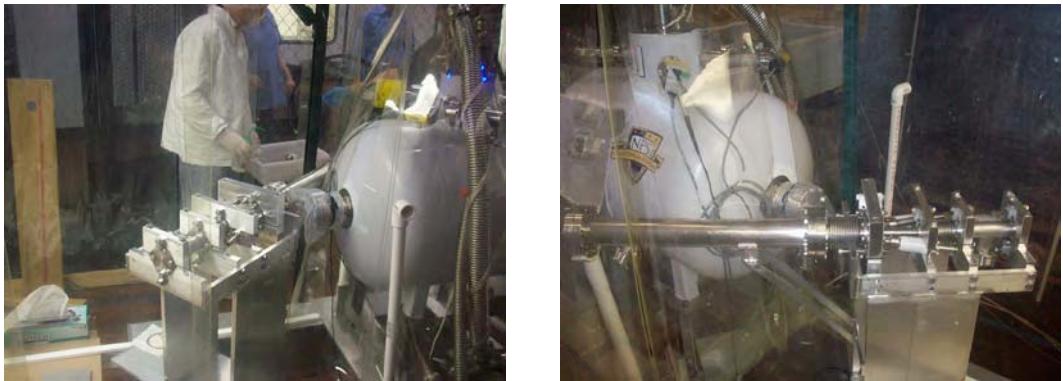


Figure 45. The cathode assembly prepared for installation into the NPS cavity. The assembly is under a portable clean room. The stainless steel surrogate beam tube and angle support are removed prior to installation.

The cathode assembly support structure is mounted upon the same foundation as the cryomodule. Aluminum pads will be placed between the transverse foot of the assembly support structure after it is in its final position, post alignment. To assist in minimizing friction during alignment and sliding the cathode assembly into the cavity, the support structure is placed on two teflon rails (similar to the ones used during the coupler insertion) with the same thickness as the aluminum pads, as

seen in Figure 46. The structure is top-heavy and extreme care must be taken during insertion to prevent the cathode assembly from striking the beam tube or interior of the cavity.



Figure 46. Two thin teflon rails are used to minimize friction when sliding the cathode assembly structure into place. The step behind the foundation is to allow the structure to slide behind the cryomodule with room for the cathode to clear the beam tube prior to insertion. One of the aluminum pads for mounting the structure to the foundation can be seen near the top of the photo, between the rails.

To prepare for removal of the surrogate beam tube, the bolts in the flange are removed using standard cleanroom procedures and cleaned. To maintain a seal between the flanges, the top and bottom bolts are reinstalled after cleaning, but only torqued to finger tight. Figure 47 shows the flanges being prepared for separation. Even though the flanges have been kept in a clean environment since removal from the clean room, the process of removing the bolts from the flanges creates particulates that could enter into the cavity once the gasket seals are broken.

Prior to opening the cavity and cathode to the environment of the portable clean room, all mating surfaces are carefully cleaned.



Figure 47. The surrogate beam tube on the cathode assembly and the blank flange on the cyromodule are prepared for removal. All surfaces are wiped down prior to opening the beam tubes to minimize intrusion of particulates generated by removing bolts from the mating flanges.

The first major issue in the assembly of the NPS gun and beamline was encountered as the surrogate beam tube was being removed from the cathode assembly. Recalling Figure 25, there is a teflon spider located approximately halfway down the cathode assembly to provide support. The spider provides a nonconducting supporting surface for the cathode assembly, decreasing the length that is cantilevered. In this case, as the surrogate beam tube was being slid along the cathode assembly

axis, the teflon spider impinged on the weld bead between the beam tube wall and the connecting flange. Significant force was necessary to remove the surrogate beam tube, resulting in deformation of the spider. Figure 48 shows a series of photos documenting this issue.



Figure 48. Top left is the initial onset of the binding of the spider. Top right is a close-up of the spider as technicians attempt to slide the surrogate beam tube off the cathode assembly. Bottom left is a picture of technicians cleaning the spider as it will be inserted into the cryomodule along with the cathode assembly. Bottom right shows the deformation of the spider after it was removed from the surrogate beam tube.

Once the teflon spider was successfully extricated from the surrogate beam tube, it was inspected to ensure it was still safe to insert into the cryomodule. As no damage other than deformation of the spider's arms was noted, it was deemed

safe to proceed. The spider was gently cleaned and the cathode assembly was moved into position for insertion into the cavity.

Figure 49 shows the cavity beam tube being opened for insertion of the cathode assembly. This figure also shows the cathode button that is installed in this assembly. The cathode button is a press-fit piece of small grain niobium whose surface was prepared using a buffered chemical polish. No further preparation of the cathode surface was performed to this point other than maintaining it in a clean environment prior to insertion into the cavity.



Figure 49. The last seal before opening the cavity beam tube is broken to insert the cathode assembly. Note the niobium cathode button is visible at the tip of the cathode assembly.

Technicians carefully maneuvered the cathode assembly along the teflon rails, ensuring the cathode stalk entered the cavity beam tube coaxially until the cathode tip was visible via a transit from the end of the beamline (described shortly), as seen in Figure 50. As the teflon spider began to enter the beam tube, the process

was again disrupted. We also see in Figure 50, the spider, due to its deformation, was now an interference fit in the beam tube, preventing the assembly from smoothly entering the cavity. Upon initial insertion, the spider, rather than entering the beam tube, merely slid along the cathode stalk.



Figure 50. On the left, technicians begin to move the cathode assembly into the cavity. On the right, the teflon spider, due to deformation, now no longer has clearance inside the beam tube to easily slide without contact. In this photo, we can see that after pulling the cathode assembly back out slightly, the spider has slid along the assembly from its intended position (where the holes are visible).

To ensure the spider was carried into the tube as near as possible to its intended position, a flexible PVC welding rod was used as a pusher. Using this method, as shown in Figure 51, one technician continually attempted to keep the teflon spider in its detente while the other slowly inserted the cathode assembly further into the cavity. It is estimated that the spider is within 2–3 cm of its intended position based on the flange gap when the pusher had to be removed. While less than ideal, this location should provide adequate support to the assembly. It is unknown but unlikely that the spider will shift position as intended when the cathode position is adjusted using the bellows.

Once the cathode assembly was inserted far enough for the flanges to mate, seen in Figure 52, the system was made vacuum tight and the final setup of the adjustment mechanism for the cathode position completed. Unfortunately, this

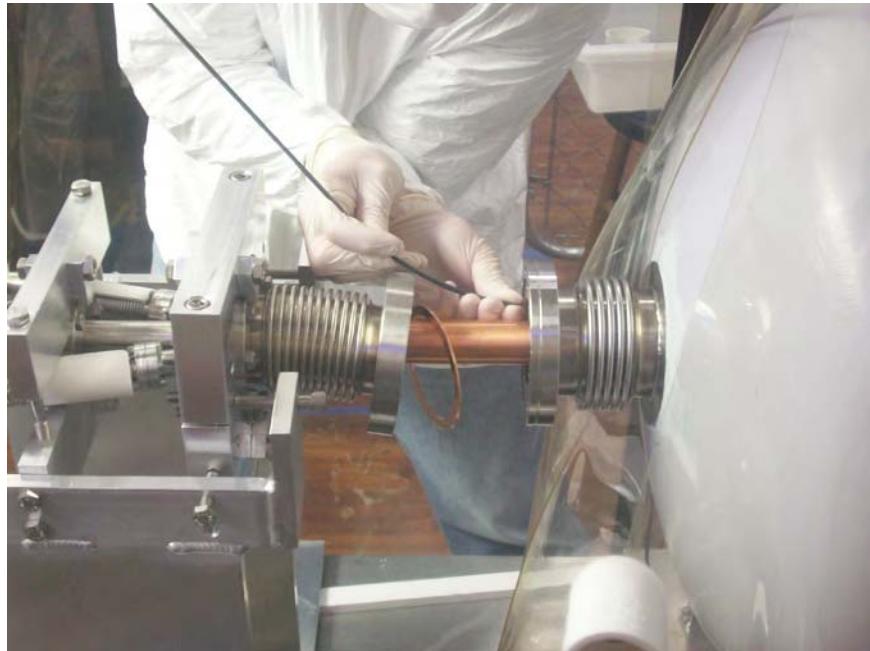


Figure 51. A PVC pusher is used to position the spider during insertion.

version of the adjustment system used custom bellow sections and both flanges at the cavity interface are nonrotatable. To obtain satisfactory bolt line up, a minimal torque has been applied to the adjustable bellows sections of the cathode assembly. This should not present a limitation in the system operation, but it should be noted that a rotatable connection to the cavity for follow-on builds would greatly ease connection of the cathode assembly to the cavity. Once under vacuum, the cavity and cathode assembly achieved vacuum in the microtorr level within an hour, indicating that all connections were leak tight and no out-gassing contaminants had been introduced.

Alignment of the cathode to the nose cone is performed using an optical transit, seen in Figure 53, through an optical window in the end of the beamline. During installation of the cathode assembly, direction from this station was provided to the technicians to ensure the cathode button was properly centered in the nose cone.



Figure 52. The cathode assembly is mated to the cavity and a vacuum tight connection is made. The cavity was then placed back under vacuum.

By providing supplemental lighting to the interior of the beam line, features that can be used for determining where the transit is focused become visible. This supplemental lighting is provided by two small flashlights, one illuminating the beam pipe just downstream of the laser mirror and the other shining through the laser port and bouncing off the laser mirror to the cavity. By adjusting the focus of the transit, one can orient oneself to the interior of the beamline as shown in Figure 54. The cathode button was visible (it had a very high contrast ratio) almost from the moment it was inserted into the beamline.

Once the cathode was in place, the drive laser needed to be aligned to the cathode. In previous work, the laser technicians had performed alignments up to the laser window on the beamline. To assist in verifying alignment, NPS provided a UV sensitive camera and UV telephoto lens to image the cathode, as seen in Figure 55. First, a helium-neon (HeNe) laser operating in the visible spectrum was used



Figure 53. Dr. John Lewellen observes the positioning of the cathode assembly relative to the cavity nose cone using an optical transit.

to align the final steering mirrors to put the HeNe spot onto the cathode surface. Once the HeNe was seen to be properly centered on the cathode, a flip-in mirror was adjusted to take the HeNe out of the final steering enclosure to a distant point and the UV drive laser aligned to shine coincident with the HeNe spot. Once the flip-in mirror was removed, both spots should shine on the same spot on the cathode with a high degree of confidence. With the HeNe laser interrupted and the focal length of the telephoto lens adjusted (it is chromatic), we were able to verify the UV laser was striking the cathode in the correct position.

After completion of the assembly and alignment activities, the NPS gun/booster is ready for RF testing and conditioning leading to an initial beam test. The various modeling and testing of the cavity will be discussed in the following chapter.

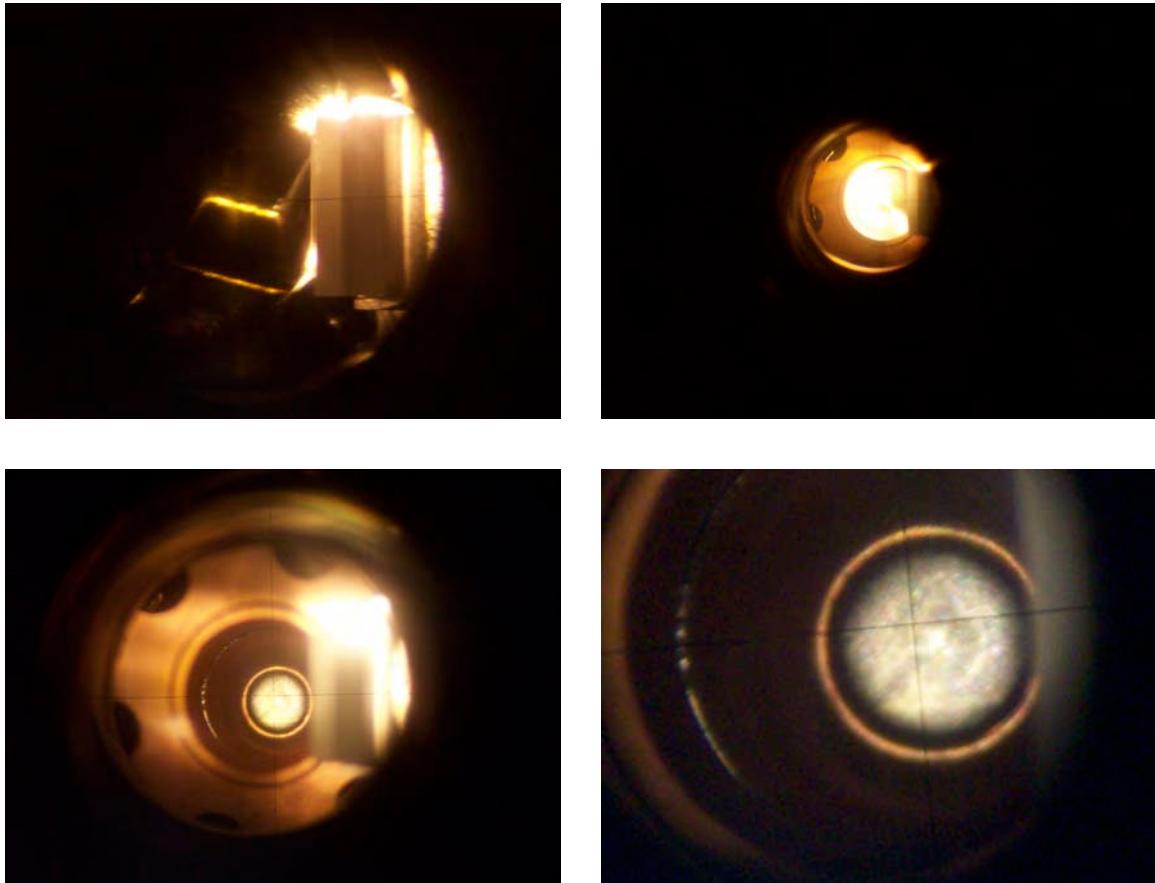


Figure 54. Changing the focus of the transit allows one to see all the interior components of the beamline or to focus on the cathode face. In the upper left is the in-vacuum laser mirror. In the upper right is the shorted end of the RF coupler. In the picture in the lower right, the grey area is the niobium cathode button, the bright ring around the button is the edge of the copper recess in which it fits, and the faint bright ring to the left of the button is the cavity nose cone.

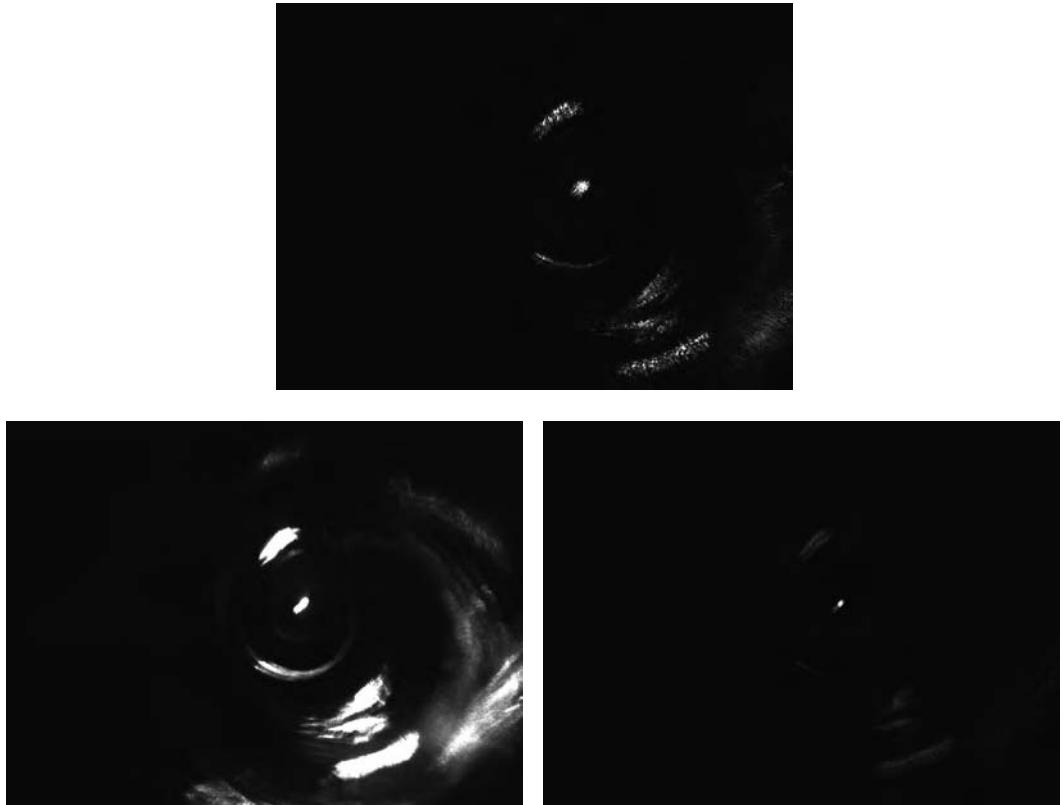


Figure 55. After aligning the cathode in the cavity, verification that the laser spot is striking the expected point on the cathode is still required. These photos are taken of both a HeNe spotting laser and the UV drive laser on the cathode. The top photo is the HeNe laser. The bottom photos are the UV laser spot with an open aperture (left) to show the multiple reflections of the scattered light and a closed aperture (right) for fine locating of the primary beam.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. MODELING AND EXPERIMENT

Modeling and experimentation are essential to any research project, especially as the size or complexity of the project increases. In the case of the NPS 500 MHz gun/booster, the prototype is a one-of-a-kind device that is also the first of its kind. There is no previous experience using the quarter-wave cavity type for the acceleration of electrons, nor for any other particle in an axial configuration. Significant time and resources were spent in developing accurate representative models and running numerous simulations to understand how the systems should react before attempting the same setup and tests on the actual hardware.

This combined approach means that during the experiment, researchers have a much better understanding of the results they should be seeing. Experimentation also provides grounding for the models, allowing them to increase their representative value to the experimental researcher. The NPS research team has followed this approach throughout the development and testing of the NPS gun/booster. Results of both simulation and experiment are presented in the chapter that follows, and where applicable, comparisons are made between the two.

A. MODELING

Modeling provides an environment where the researcher is able to bring almost any variable into the design space. It can allow researchers to examine parameter spaces beyond what their experimental environment allows. On the other hand, it is possible to develop a model that violates or fails to incorporate fundamental rules such that simulations no longer have a basis for comparison to experiment. The researcher must be extremely careful to ensure his models are grounded to the experiment of interest and that simulations do not exceed the capabilities of the programming developed to run them.

1. Simulation Codes Used

Numerous simulation and analysis software packages were and continue to be used in the design and development of the NPS superconducting radiofrequency (SRF) gun/booster. During the design phase of the cavity, the majority of the simulation work was performed using the Poisson Superfish family of codes developed by the Los Alamos Accelerator Code Group.¹ Particle interactions with the cavity have been modeled in Parmela and General Particle Tracer (GPT) and benchmarked occasionally against other accelerator codes. We have also occasionally used new software packages to evaluate their usefulness by comparing their results to data already obtained in previous simulation work on this cavity. We will primarily focus on Poisson Superfish, Parmela, and GPT as these software packages were used for the bulk of the work, but will mention other codes used for benchmarking where appropriate.

Sub-applications within each code for massaging of input decks or external data manipulation for analysis will not be discussed as they are not integral to understanding the simulations performed. Additional applications, such as the SDDS (Self-Describing Data Set) toolkit from Argonne National Labs,² MATLAB, and Microsoft Excel were used to sort, manipulate, and present data at various times throughout this research. Optimization and parameter searches were either performed by simple self-generated scripts or through the use of Tcl³ (a cross-platform scripting language) optimizer and sequencer scripts provided by Dr. John Lewellen of the Naval Post-graduate School.

All simulations were performed on Intel-based Apple Macintosh computers. Large simulations were handled using a Mac Pro with two 2.8 GHz Quad-Core Intel Xeon processors and 32 GB of RAM. Where necessary to use a Windows-based

¹Superfish documentation and software can be downloaded from the LAACG at http://laacg1.lanl.gov/laacg/services/download_sf.phtml.

²The latest versions may be downloaded from Argonne National Lab at http://www.aps.anl.gov/Accelerator_Systems_Division/Operations_Analysis/software.shtml.

³ActiveState at <http://www.activestate.com/activetcl> provided the free Tcl installation used in this work.

environment, Parallels version 5 was used to host a Windows XP Pro 64 bit environment with access to 6 processors and 8 GB of RAM. Additional small simulations and parameter searches were performed on a MacBook Pro with a 2 GHz Intel Core Duo processor and 2 GB of RAM. Parallels version 5 hosting a Windows XP Pro 32 bit environment with one processor and 1 GB of RAM was used where necessary. Simulation comparisons between the two computers were excellent except in the time required to complete the computation.

a. Poisson Superfish

Poisson Superfish is a family of codes generally used for modeling of electromagnetic systems [30]. Poisson is typically used for electro- and magneto-static models, while Superfish is designed to handle RF electromagnetic fields. Both of these solver programs use either two-dimensional Cartesian or axially symmetric cylindrical coordinates. Poisson was used to model the magnetic fields of the superconducting solenoid and Superfish was used to model all RF cavities examined in this work. Other elements, including a normal conducting solenoid, window-frame beam steering magnets, and quadrupoles were modeled in Poisson as well. Since Poisson and Superfish solve different problems, we discuss their approaches separately.

Poisson, the static field solver, takes a text input file describing the geometry of a problem and the relevant properties such as boundary conditions and materials and solves for the vector potential, \mathbf{A} . The first step of the algorithm is to mesh the area of interest using an irregular triangular mesh. Each triangle touching a boundary will have two vertices touching the boundary, altering the boundary shape to line segments approximating the intended curve. Poisson then solves Ampere's Law at each mesh point,

$$\oint_{\mathcal{C}} \mathbf{H}(\mathbf{r}) \cdot d\mathbf{l} = \oint_{\mathcal{C}} \left(\frac{\nabla \times \mathbf{A}}{\mu_0} - \mathbf{M} \right) d\mathbf{l} = \int_{\mathcal{S}} \mathbf{J}(\mathbf{r}) \cdot \hat{\mathbf{n}} dA , \quad (\text{IV.1})$$

where \mathcal{C} is the contour of integration, \mathbf{H} is the auxiliary magnetic field, \mathbf{M} is the

magnetization, \mathbf{J} is the current density, \mathcal{S} is a surface of area A enclosed by \mathcal{C} , and $\hat{\mathbf{n}}$ is the outward normal. It approximates that over each triangular area the vector potential, \mathbf{A} , is linear, the reluctivity is constant, and the current density is constant. From \mathbf{H} , the program can solve for \mathbf{B} , the magnetic field by

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}) . \quad (\text{IV.2})$$

Additionally, the program uses the Coulomb gauge,

$$\nabla \cdot \mathbf{A}(\mathbf{r}) = 0 . \quad (\text{IV.3})$$

The specifics of the algorithm for solving equation (IV.1) can be found in [31], but in general, for each mesh point the program solves a contour integral through the surrounding six triangles. The value for the mesh point is updated and the process continues through the entire mesh iteratively until a convergence criterion for the vector potential is satisfied. Boundary conditions are applied as either Dirichlet (the value of the vector potential is given on the boundary) or Neumann (the normal derivative value of the vector potential on the boundary is given) boundary conditions.

These conditions are specified in the input file.

Superfish also solves two dimensional cartesian (assuming $\hat{\mathbf{z}}$ dimension is infinite) and three dimensional cylindrically symmetric problems. It, however, focuses on solving RF problems. Primarily, Superfish solves for resonant frequencies of cavities and waveguides and then determines the electromagnetic fields resulting from the problem geometry and frequencies of interest.

Superfish starts from the generalized Helmholtz equation

$$\nabla^2 \phi + k^2 \phi = S , \quad (\text{IV.4})$$

where $\phi(x)$ are the eigenfunctions and S is a source term. The modes (TE or TM) of a cavity are determined from the boundary conditions specified by the user. The geometry of the problem is meshed exactly the same as for Poisson. The constitutive

relations for the assumed homogeneous, isotropic, non-conducting medium in the cavity are

$$\mathbf{D} = \epsilon \mathbf{E} \quad (\text{IV.5})$$

$$\mathbf{B} = \mu \mathbf{H} , \quad (\text{IV.6})$$

where ϵ is the permittivity and μ is the permeability as input by the user in the problem definition.

For solutions presented in this work, a field drive point is provided by the user, which allows Superfish to use the auxiliary magnetic field as the eigenfunctions,

$$\nabla \times (\nabla \times \mathbf{H}) - k^2 \mathbf{H} = 0 . \quad (\text{IV.7})$$

Applying Stokes Theorem, we get the integral equations

$$\int_{\mathcal{S}} \nabla \times (\nabla \times \mathbf{H}) \cdot \hat{\mathbf{n}} \, dA = \oint_{\mathcal{C}} (\nabla \times \mathbf{H}) \cdot \, d\mathbf{l} = k^2 \int_{\mathcal{S}} \mathbf{H} \cdot \hat{\mathbf{n}} \, dA , \quad (\text{IV.8})$$

where \mathcal{S} is the area bounded by the contour \mathcal{C} and $\hat{\mathbf{n}}$ is the outward normal of the surfaces. The same process of integrating around the mesh triangles surrounding the point of interest is implemented in the Superfish solver and iterating until the convergence criteria is met.

To find resonant frequencies, Superfish assumes a fictitious magnetic current density \mathbf{K} , which is used to drive the fields in the cavity. At resonance, the current required to drive the fields goes to zero. The function

$$D(k^2) = k c \frac{\int \mathbf{H} \cdot \mathbf{K} \, dV}{\int \epsilon (\mathbf{H} \cdot \mathbf{H}) \, dV} \quad (\text{IV.9})$$

is used to search for zero crossings. A zero crossing is necessary but not sufficient for determining a resonance as

$$\frac{dD(k^2)}{d(k^2)} = -1 \quad (\text{IV.10})$$

must also be true for a particular crossing to be a resonance.

When the user provides a frequency of interest, Superfish will solve for the fields at a resonance near this frequency. Subroutines can also be used to find many pertinent values of interest including energy per unit volume in the cavity, power losses on the cavity walls, average accelerating field, shunt impedance, cavity quality factor, maximum field values along boundaries, and many other parameters based on the user’s input assumptions. Of critical importance is the ability to output field profiles and maps that can be imported into other codes, such as Parmela and GPT for use in modeling particle trajectories.

b. Parmela

Parmela is a particle accelerator design and simulation code developed at Los Alamos National Lab beginning in the 1980s. As with Poisson Superfish, the code is maintained and distributed by the Los Alamos Accelerator Code Group.⁴ Parmela uses an input file that describes the accelerator geometry, fields, particles, and related subroutines to be invoked. The internal library of accelerator equipment including cavities and correctors is extensive, but arbitrary fields may also be imported. Particle distributions of multiple species and charge signs can be invoked simultaneously and distributions may be generated within Parmela or imported from an external source. The code tracks particles through the accelerator based upon a phase angle clock based on a designated RF frequency rather than an explicit time reference. Particle transport is handled by various subroutines based on the dynamics invoked in the input file. The number of particles in a simulation is limited only by the available memory of the computer performing the simulation. Each specified particle is really a macroparticle representing many charged particles, but maintaining the appropriate charge to mass ratio for particle dynamics.

RF field conventions in Parmela are based upon a “sine” convention in which $\phi = 0$ indicates a zero crossing. Since Parmela assumes a particle charge of

⁴Parmela may be downloaded, after approval, from the LAACG at <http://laacg1.lanl.gov/laacg/services/services.phtml>

+1 unless designated otherwise, the building RF fields from this point represent an accelerating field in the positive \hat{z} direction. There is no issue assuming a +1 charge for electrons, so long as the convention is maintained throughout (i.e. reversing fields for bending magnets).

Particle tracking is handled through an internal database that compares a particle’s position with the beamline geometry. When particles are lost due to interaction with the beam line geometry or beam line elements, the particle’s database entry is annotated as “lost” and is no longer tracked.

Particle motion is imparted by the code based on local and global field elements and interactions with other particles. At each phase step (or specified number of steps), each particle receives an impulse resulting from space charge, fields from the local element, and distributed fields. Space charge fields are probably the most difficult to understand, and Parmela has many implementations to consider. The two primary forms are 2D and 3D. The 2D form is used primarily for axisymmetric beams with particles distributed upon a 2D mesh and Poisson’s equation solved for space charge forces. In the 3D form, a mesh is formed around the bunch and used for calculating particle-particle interactions. In all cases, Parmela transforms to the moving frame of the bunch to perform the particle-particle space charge interactions.

Much more detailed information can be found in the Parmela users manual [32]. There is significantly more detail required to learn to use the program efficiently, it is only our intent to provide a general understanding of the program sufficient to understand the results presented. Data output is to a binary data file for use by Parmela’s dedicated interpreter program. For this work, a translator was used to transfer this binary data to SDDS format for use with previously developed analysis and manipulation tools. Parmela is a Windows command line program only.

c. General Particle Tracer

General Particle Tracer (GPT) is similar to Parmela in that it tracks particles through input fields and geometries, but has a more generalized approach as

its use is not intended to be specifically limited to particle accelerator studies. GPT also uses the macroparticle approach identical to Parmela. Many of the features and interactions with the executable code are similar in that a geometry and fields are specified from the internal libraries or imported from external sources. Multiple species and charges may be handled simultaneously. Particle distributions may be generated internally or imported. Particle transport is solved in the time domain by a 5th order embedded Runge-Kutta integrator with adaptive step size control [33].

Data output is to a binary database that can be directly interpreted by the Windows GUI environment or translated to text, SDDS format, or other specific formats via included translators. Data acquisition is initiated by user-stipulated times or positions. In time-specified output, all particle data is collected at each specified time. For positional output, non-destructive “screens” are placed at the required position and particles recorded as they pass through the desired location. GPT will interpolate to determine the coordinates and time when the particle passed through the desired screen location.

Numerous space charge routines are available for use, including 2D and 3D models similar to Parmela’s. The two routines predominantly invoked in this work are the `spacecharge3D()` and `spacecharge3Dmesh()` routines. The mesh version calculates space charge forces in the rest frame of the bunch approximating the charge density at the nodes of a non-equidistant mesh encompassing the bunch. The charge density is then fed into a Poisson solver to obtain the electric potential, which is differentiated to obtain the rest frame electric fields. These fields are combined with the system geometry fields and applied as impulses to the particles. This particular routine was developed and optimized with high charge and high brightness bunches in mind [34]. The `spacecharge3D()` routine is a particle-particle solver using the fully relativistic field equations. It makes no approximations, but has a significant cost in CPU time over the meshed approach [33]. This routine was used only when modeling low energy beams or when very few particles were being modeled.

Of particular interest is that GPT is fully customizable. If a particular beam line element is used repeatedly, it is possible to write a custom element that can be compiled into the GPT program for use. GPT can be executed in a Windows graphical user interface, Windows command line, and Linux environments.

2. Beam Simulation Methodology

To ensure comparability between beam simulations using different software packages and different component models, a standard methodology was employed. In general, it is very difficult to perform multidimensional parameterization by hand. To simplify this, beam lines were approached in sections, limiting the number of free parameters available for analysis and adjustment and terminating the propagation at specific points associated with the sections to limit processing time. Parameters controlled by sections previous to the section of interest are held fixed and an initial parameter search through the current chunk is performed. After analysis of the resulting data, a carefully chosen starting point and limited parameter space is provided to the optimizer code. The optimizer attempts to adjust the free parameters it is given to minimize a figure of merit (FOM) provided by the user. The resulting optimized variables are incorporated and the next section is taken into consideration.

As an example, let us consider optimization of beam radius in the context of the 500 MHz gun. At the start, a cathode position, bunch charge, beam shape, etc., have already been determined. The first section is the accelerating cavity itself (terminating prior to the next major beam line element—in this case, the superconducting solenoid) and the variable of interest is the cavity launch phase. For a FOM, we desire to maximize the average energy of the beam, minimize the spot size at the termination, and transport the maximum charge, with transport being more important than energy or spot size. To account for this, we set up the FOM,

$$FOM = \frac{\sigma_r}{E \left(\frac{N_{out}}{N_{in}} \right)^5}, \quad (IV.11)$$

where σ_r is the beam radius, E is the average kinetic energy of the beam, N_{in} is the

number of particles in the beam at instantiation, and N_{out} is the number of particles at the last data recording location. Note that the transmission of particles is raised to the fifth power to account for the severe penalty applied for lost particles. The number of particles chosen is dependent on the task being performed—in general, a parameter search is performed with fewer particles than an optimization run since many more runs are typically performed for a parameter search.

The output data from this first section is analyzed and a range of viable phases is provided to the optimizer with an estimated best value and an initial step size. The optimizer then looks for a minima in the FOM by adjusting the allowed free parameters. Once an optimal launch phase is found, that value is held fixed for the next chunk. In this example, that would be propagation from the superconducting solenoid to the normal conducting solenoid.

For the second section, the applied magnetic field of the solenoid is allowed to vary to optimize transport and spot size (meaning the system will attempt to focus at the final data recording point). The same parameter search process is applied to the new variable and an optimal magnetic field determined. The final section is to include the normal conducting solenoid and terminate at some down stream point (taken in the following simulations to be the assumed location of the next beam line element).

Once parameter values have been determined through the entire beamline, these are used as starting values for an entire system optimization. In general, the values found in the above process provide good estimates of the optimized values, however some modification or interaction by the user is necessary. For instance, attempting to determine the optimal focusing of a solenoid prior to a booster cavity requires looking at both the magnetic field strength of the solenoid and the cavity phase simultaneously, as there can be multiple local minima in the resulting parameter surfaces. This requires external bounding to ensure the optimization routine remains

near an optimal solution. Where deviations from this general process occurred, they are described in the simulation description.

3. Space Charge

The primary force making transport difficult in high bunch charge, low energy beams is the self-force of the electrons in the beam interacting with each other. This is normally termed “space charge.” As one attempts to push the charges closer together, creating smaller/shorter bunches, the forces acting to expand the beam increase significantly. For high power FEL applications, it is desirable to push the peak current of the beam as high as possible.⁵ This requires a beam with high-bunch charge, and short temporal extent. Since this concept is so important to simulations performed in this research, it is worthwhile to review some of the equations and some simplified models to develop some instinct into how space charge affects beams of interest.

We can use continuous beam theory as a zeroth order approximation to the moderate bunch lengths possible with the current NPS drive laser, as the bunch length is long compared to the gaps between the electrodes in the Stanford gun. Additionally, insight developed here will be useful when our focus shifts to RF guns. We start with a simplified derivation of the envelope equation [4] and compare its results with a simple simulation that includes longitudinal effects. To start the derivation, we assume a uniform charge density, laminar (no large divergence), steady state current with the geometry as shown in Figure 56.

This current has two self fields in the lab frame—a radial electric field described by Gauss’ law and a magnetic field described by Ampere’s law in SI units,

$$E_r = \frac{I}{2\pi\epsilon_0 v} \frac{r}{R^2} \quad (\text{IV.12})$$

$$B_\theta = \frac{\mu_0 I}{2\pi} \frac{r}{R^2} . \quad (\text{IV.13})$$

⁵FEL theory explaining this will be detailed in the next chapter.

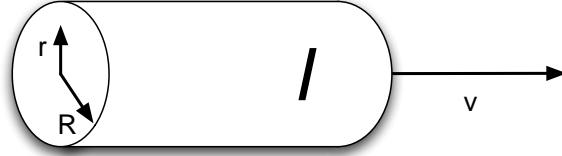


Figure 56. A current distribution of radius, R , moving with velocity, v , having current, I .

Applying the relativistic Lorentz force, we find that the radius of the beam changes over time as

$$\begin{aligned}
 \frac{d^2r}{dt^2} &= \frac{q}{\gamma m} (\mathbf{E} + \mathbf{v} \times \mathbf{B})_r \\
 &= \frac{qI}{2\pi m \gamma} \frac{r}{R^2} \left(\frac{1}{\epsilon_0 c \beta} - c \beta \mu_0 \right) \text{ using } v = c \beta \\
 &= \frac{qI}{2\pi \epsilon_0 m c \beta \gamma} \frac{r}{R^2} (1 - \beta^2) \\
 &= \frac{qI}{2\pi \epsilon_0 m c \beta \gamma^3} \frac{r}{R^2} .
 \end{aligned} \tag{IV.14}$$

Changing from time to propagation distance, $z = c\beta t$, we rewrite equation (IV.14) as

$$\frac{d^2r}{dz^2} = \frac{qI}{2\pi \epsilon_0 m c^3 \beta^3 \gamma^3} \frac{r}{R^2} . \tag{IV.15}$$

The leading coefficient can be defined as the generalized perveance,

$$\mathcal{K} \equiv \frac{qI}{2\pi \epsilon_0 m c^3 \beta^3 \gamma^3} ,$$

which is a measure of the space charge forces present in a relativistic charged particle beam.

If we now include a driving term (such as from a solenoid field) to equation (IV.14) with coefficient k_0^2 proportional to the particle beam radius and an emittance term, we are very close to the envelope equation. To get to the transverse envelope equation, we only consider how a particle at the outer edge of the beam will behave, and we have

$$R'' + k_0^2 R - \frac{\mathcal{K}}{R} - \frac{\epsilon_n^2}{R^3} = 0 . \tag{IV.16}$$

The normalized emittance, ϵ_n , can be viewed as a measure of the internal disorder of the beam. A higher emittance beam will diverge faster than a lower emittance beam. In particular, the normalized emittance is a statistical quantity of trace space that describes the area encompassed by the group of particles making up the beam. Mathematically, the normalized emittance is defined

$$\epsilon_n \equiv 4\beta\gamma\sqrt{\langle x^2 \rangle \langle (x')^2 \rangle - \langle xx' \rangle^2}, \quad (\text{IV.17})$$

where x is the transverse coordinate displacement from the axis and x' is the transverse velocity of a particle.

The envelope equation gives the researcher insight into how the maximum extent of a well-defined, long beam behaves under space charge forces. But with the simulation tools available, we can also look at some more representative beams and how they behave as compared to the envelope approach. To start, we define bunch geometries of interest that will be used in further simulations in this work. Figures 57, 58, and 59 show pertinent data for three basic beam shapes, a uniform distribution—uniform distribution longitudinally and radially, a Gaussian distribution—Gaussian longitudinally and radially, and an asymmetric shape—a fast rise and slower fall longitudinally and Gaussian radially. To model the envelope equation, we use a simple MATLAB script, setting the emittance to zero so that radial growth is due solely to space charge forces.

In Figure 60, we see three beam comparisons. The solid lines are best representative of the assumptions made when we derived the envelope equation—that the beam is long compared to its radius. As the beam is shortened significantly in the longitudinal direction as compared to the radial direction, the envelope equation significantly overestimates the radial beam expansion. The same value of charge is used for each calculation, however the GPT simulations take into account the increased longitudinal forces in the “pancake” beam case, quickly decreasing the charge density with a resulting decrease in radial beam expansion. Additionally, we see between the left and right plots the effects of the γ^{-3} dependence in the space charge term as

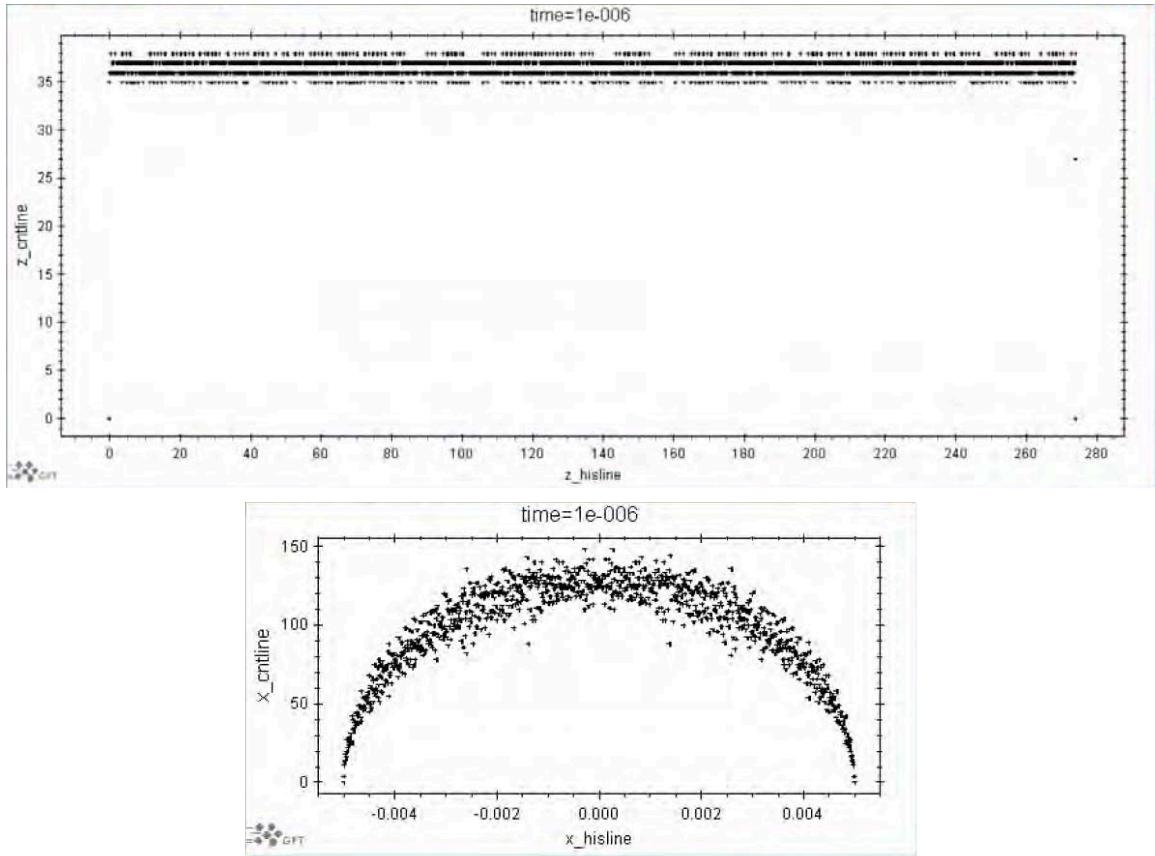


Figure 57. A distribution that is longitudinally and radially uniform. The histograms show particle distributions in z (upper) and x (lower). The distribution is shown after generation using GPT.

there is almost a two order of magnitude reduction in beam expansion of the envelope equation estimates with a two order of magnitude increase in energy (KE goes from 10 keV \rightarrow 1 MeV, γ^{-3} goes from 0.942 \rightarrow 0.0387).

The longitudinal forces are important as photoinjectors have become fairly common as electron sources to linear accelerators. Even with the 500 MHz operating frequency of the NPS cavity, a 40 ps pulse represents 7.2 degrees of phase for the cavity fields. As the cavity frequency increases, the pulse length (in phase) must decrease to avoid energy spread in the beam. Ultrashort pulses with high bunch

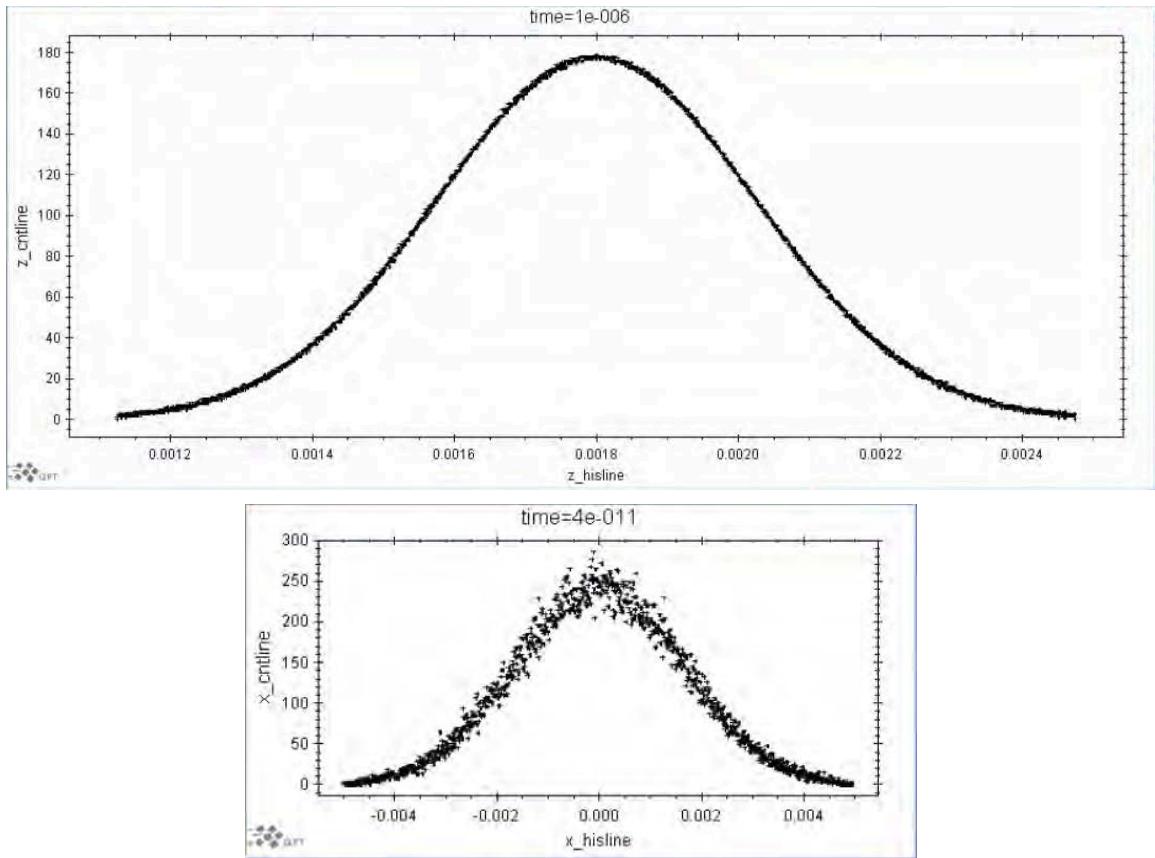


Figure 58. A distribution that is longitudinally and radially Gaussian. The histograms show particle distributions in z (upper) and x (lower). The distribution is shown after generation using GPT.

charges have been shown to lead to elliptical electron bunches having extremely desirable propagation characteristics [35], thus, being able to generate these type pulses is desirable.

Knowing the longitudinal expansion of the pulse is important, and in the short pulse length regime, possibly desirable for propagation characteristics. We compare the three pulse types with that of the long pulse generated envelope equation. Using a common beam energy of 100 MeV, a bunch is started from the origin and allowed

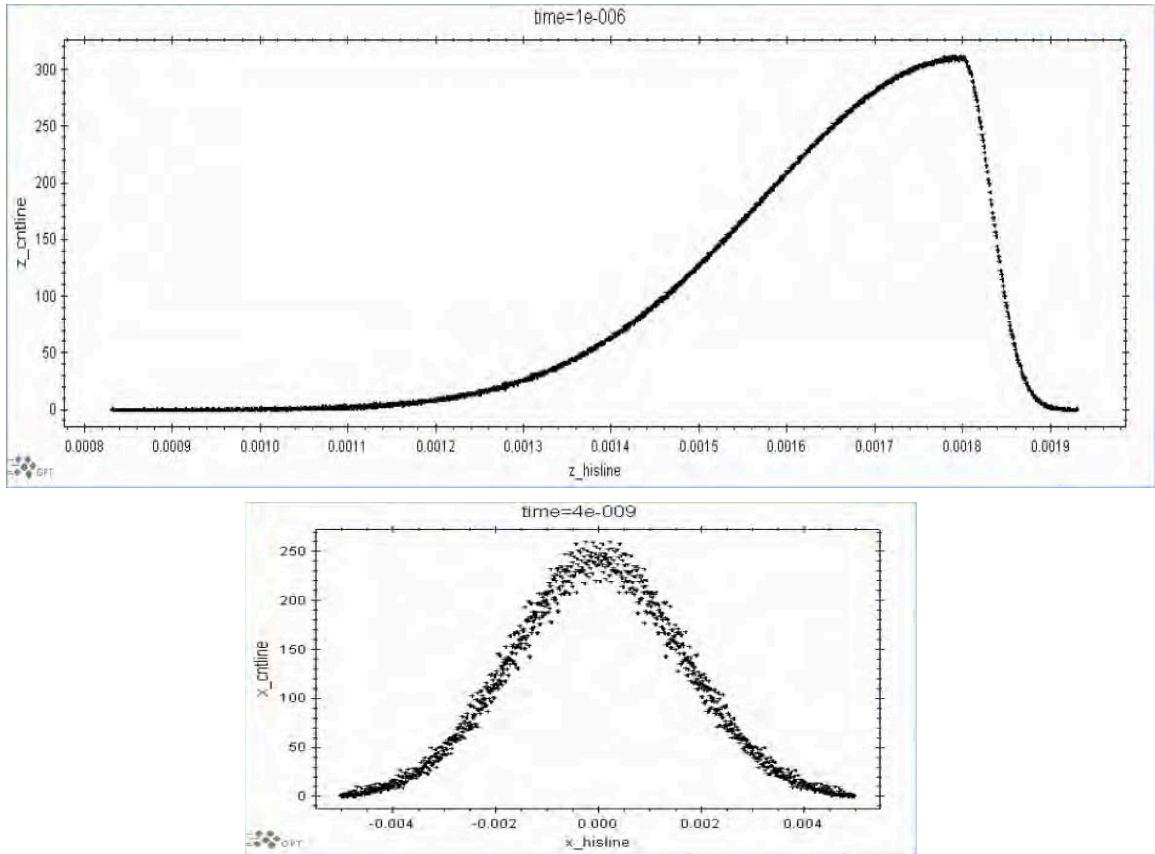


Figure 59. A distribution with a fast rise and slow fall longitudinally and a Gaussian shape radially. The histograms show particle distributions in z (upper) and x (lower). The distribution is shown after generation using GPT.

to propagate in free-space under only the influence of internal space charge forces. Table 4 shows the resulting maximal radial, σ_r , and longitudinal, σ_z , dimensions of the beams as simulated in GPT.

As expected, the long, thin beam expanded significantly less than the other higher charge density beams. What is surprising is how little difference there is between the Gaussian and asymmetric shapes. They both have central, or near central, high density areas that quickly expand under space charge forces. Accelerated radial growth is slowed by the rapid reduction of charge density as the short pulses “blow out.” When compared to a long pulse with an equivalent charge density (using

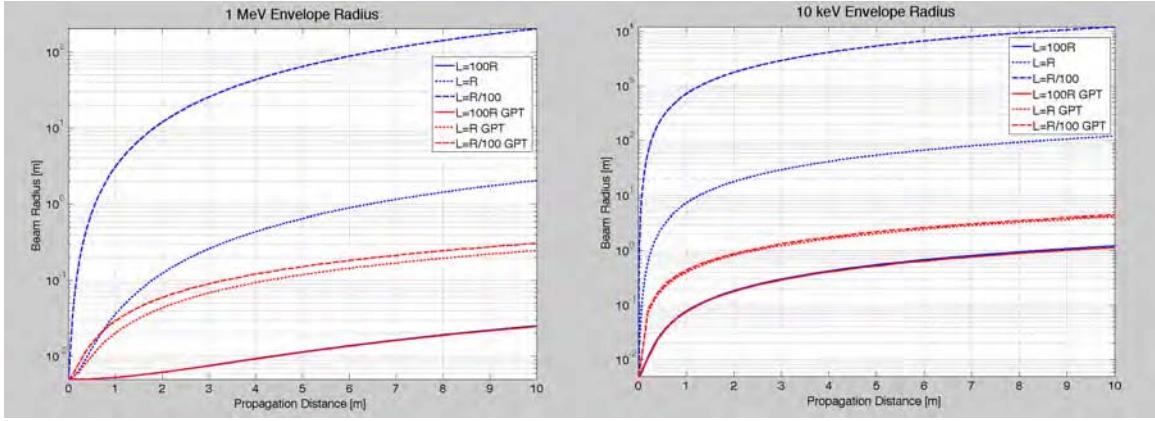


Figure 60. Space charge effects on different aspect ratio beams at low and high energies. Blue plots are derived from the envelope equation and red plots are results from GPT simulations. The plot on the left is at 1 MeV energy, and the plot on the right is at 10 keV energy. A uniform bunch charge distribution is used.

the short, uniform pulse as a basis) of ~ 1.1 pC/mm³, we see a radial expansion beyond that of the short pulse. This is probably due to the fact that the charge in the middle of the beam, where the radius increases the most, cannot decrease the bunch charge density as quickly due to the longitudinal extent of the beam preventing the pulse's longitudinal expansion.

It is apparent from our studies of space charge that external forces are required to keep the electron pulses we generate from expanding significantly. For FEL purposes, the peak current is of critical importance to the FEL interaction. As the beam expands radially and longitudinally, not only does transport become more difficult (due to nonlinearities in the focusing and bending element fields), but the peak current decreases as well (due to having the same charge in a longer, wider electron pulse). External forces for transport, such as solenoids and quadrupoles, to refocus the beam are required, especially when the beam is at low energy and space charge forces are significantly more important. Additionally, we can mitigate some of the longitudinal expansion of the beam by accelerating to higher energy as soon as possible, thus increasing γ . If one can induce a negative energy chirp (the head of the

Table 4. Many photoinjectors operate in a “pancake” regime where the electron pulse is short compared to its radius. Comparisons are presented between a 1 nC long pulse with length = 100 R and 1 nC, 40 ps pulses of various pulse shapes at a kinetic energy of 100 MeV. All “final” measurements are taken after propagating 100 m.

Beam Type	σ_r Start	σ_r Final	σ_z Start	σ_z Final
Long Thin	0.005 m	0.436 m	0.50 m	0.62 m
Short Uniform	0.005 m	2.23 m	0.0113 m	1.14 m
Short Gaussian	0.005 m	2.99 m	0.0085 m	1.48 m
Short Asymmetric	0.005 m	3.16 m	0.0072 m	1.85 m
Long, high charge density	0.005 m	3.60 m	0.50 m	1.49 m

beam at lower energy than the tail), a dispersive section, such as a chicane, can be used to decrease the pulse length, increasing the peak current.

4. DC Gun Analysis

The first experimental system that was modeled during the process of developing an injector section for the future NPS FEL was the Stanford DC gun. Initially, the plan was to use the Stanford gun as the electron source with the SRF 500 MHz cavity as a booster to bring the energy to near 1.5 MeV prior to injection into the main linear accelerator. As part of the experimentation process, models were developed for the gun and associated diagnostic beam line by the author in preparation for experimental verification.

a. *Stanford DC Gun*

The former Stanford DC gun was initially installed as an electron source for a small diagnostic beam line. The hope of the experiment was to develop experience running high voltage systems as well as testing the viability of the gun for use as the NPS FEL gun. The electron source is a gridded thermionic cathode embedded within a static voltage accelerating structure. Recalling Figure 6, the cathode sits at the electrical axis of the gun surrounded by a Pierce electrode to provide initial focusing as the beam is launched [4]. The three electrodes are electrically isolated. The Pierce electrode and cathode are maintained at large negative potential and the final electrode is at ground. The ratio of voltages between the Pierce electrode and focusing electrode is step-wise variable by choosing its potential from steps on a voltage divider seen in Figure 61.

In operation at Stanford, the DC gun was used in thermionic mode. By energizing a resistive heater behind the cathode surface, the cathode temperature was increased ($T > 1000$ K) such that electrons were “boiled” off the surface continuously. This is an extremely reliable method for generating high currents, but the current is continuous and thus not ideal for RF accelerator applications. To pulse the beam a secondary field is generated between the grid and the cathode surface opposing the accelerating fields in the gun structure. Since the grid is much closer to the cathode surface, the voltage on the grid can be less than 100 V and hold off the gun



Figure 61. The DC gun voltage divider showing the resistive ladder on the right connecting ground to the cathode potential. The white wire coming off to the left is an adjustable connection to the focusing electrode.

fields. By switching the grid voltage, the electron beam can be pulsed. Unfortunately, the switching cannot be performed arbitrarily fast, so generating temporally short, high charge bunches requires additional tools. In the case of the Stanford FEL, a sub-harmonic buncher cavity was used—the beam passes through the cavity during a field zero crossing such that the head of the beam is decelerated and the tail is accelerated compressing the beam longitudinally.

For NPS purposes, we desired to use the gun in a combined photo-thermal mode: heating the cathode to increase the electron kinetic energy, but not to the point of liberating them from the cathode surface, and providing the liberating energy via a drive laser [17], [18]. The drive laser also acts as the gating mechanism,

allowing for much higher repetition rates and shorter pulse durations. Initial experiments indicate this approach is sound, and “first beam” in the NPS Beam Physics Lab facility was achieved 17 April 2009 as captured in the inset of Figure 62. The drive laser for these experiments was the first generation version of the minimal drive laser design presented earlier. From this system, the first attempts at generating beam showed significant discrepancies between the expected charge liberated and the charge collected at the dump.

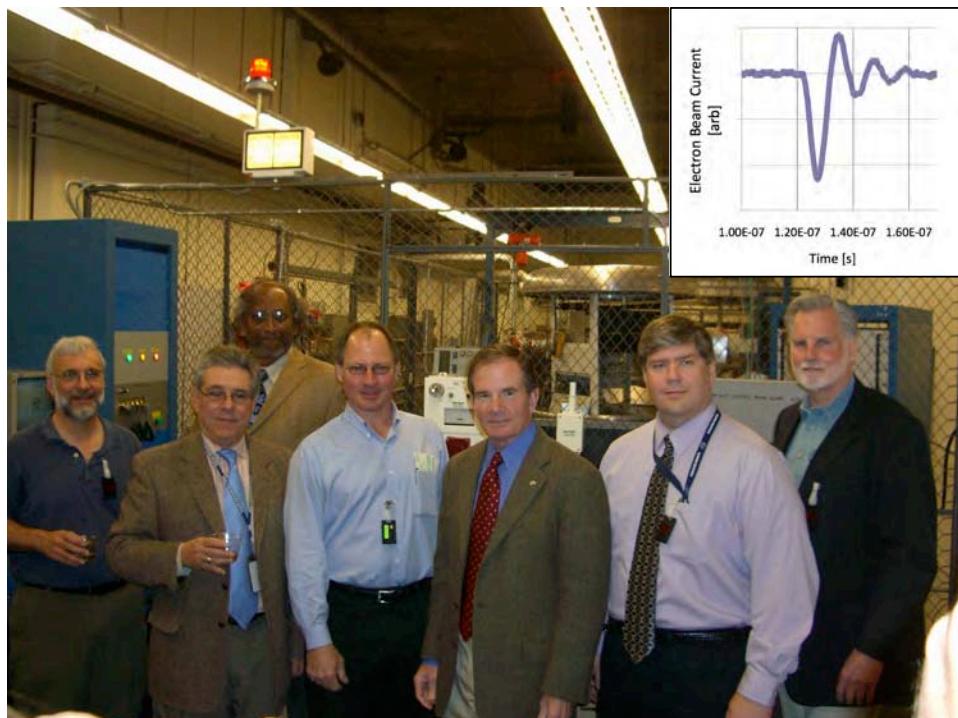


Figure 62. NPS staff and researchers celebrate “first beam” of the NPS Beam Physics Lab, 17 April 2009. In the inset is a plot of the current captured from this first measurement.

The NPS DC gun beamline, shown in Figure 63, consists of a DC gun, a four-way cross for laser entry and mirror insertion, a solenoid for electron beam focusing, a large six-way cross for vacuum pumping, and an instrument section with Faraday cups and a residual gas analyzer. The DC gun can operate with potentials

up to 250 kV, provided by a high voltage power supply. The Faraday cup section allows for crude beam centering as the two metal blocks are electrically isolated and configured similar to Figure 64, with both plates being connected to an oscilloscope for integration of the two signals.

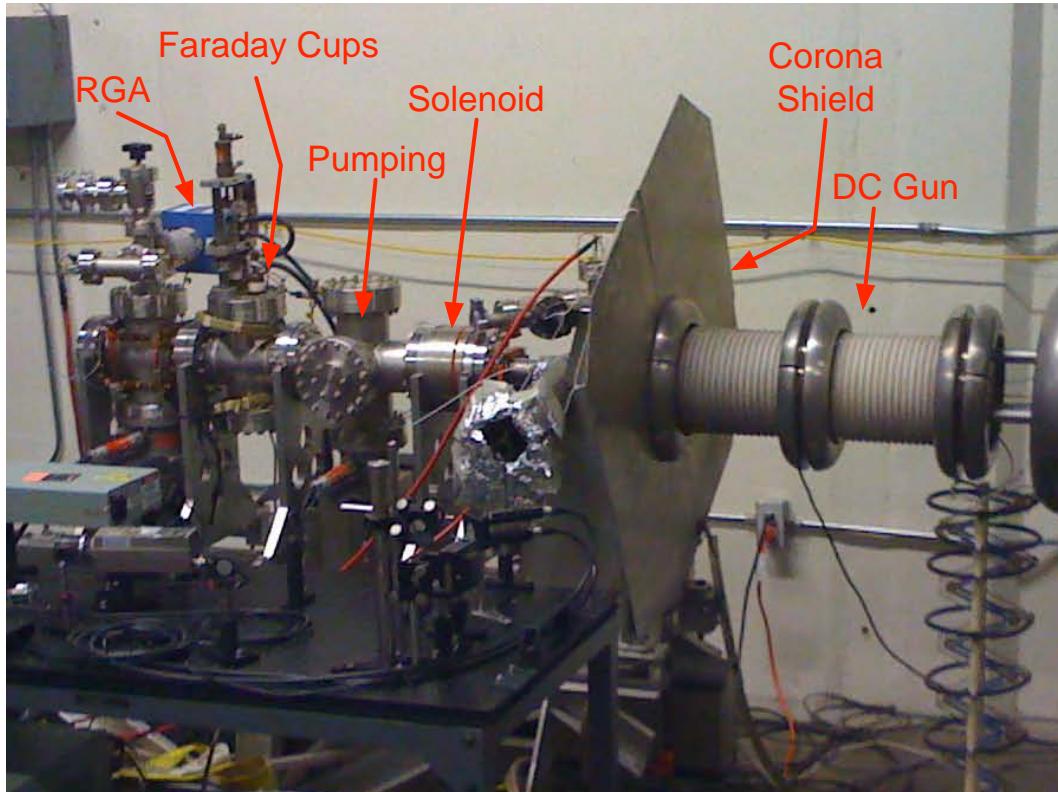


Figure 63. The DC gun diagnostic line for photo-thermal electron beam generation.

Using the best drawings available (the drawings provided with the Stanford gun show a third stage that was apparently never built), the internal dimensions and geometries were input in Poisson and solved for static electric fields. The desired operating practice is to hold the Pierce electrode and cathode at -240 kV, the focusing electrode at -140 kV, and the final electrode at ground potential. The geometry and resulting equipotential lines are shown in Figure 65. In actual operation, we were unable to condition the gun to its design voltage. However, because the voltage divider

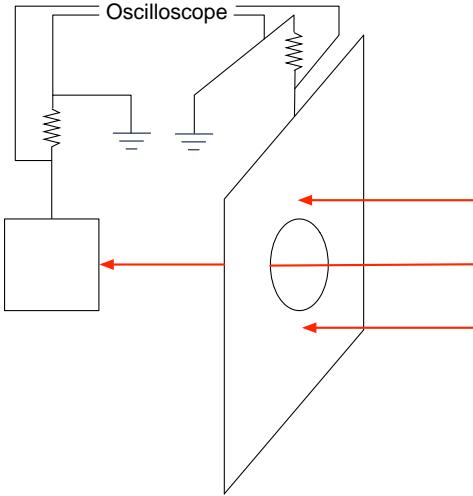


Figure 64. The Faraday cup design in the diagnostic line also allows for determination of beam centering during transport. The first Faraday cup consists of a copper plate with a central hole approximately 1 cm in diameter. The second Faraday cup is a copper block approximately 3 cm in diameter. Prior to taking measurements, the voltage signal on the second Faraday cup is maximized and the voltage signal from the first Faraday cup is minimized, ensuring the beam is properly centered as it passes into the diagnostic station.

only allows for stepwise settings between the electrodes, the ratio of cathode/focusing voltages remain constant regardless of the operating voltage at the time. By scaling the field map in the simulation software, we can model any potential value for a given electrode voltage ratio.

The initial diagnostic beam line was planned for developing procedures and initial measurements of quantum efficiency and bunch charge. Using the author's experience in naval engineering, these were developed to be easily understood and transcribed to fleet language. Additionally, the experience developed in these experiments is directly applicable to experiments planned for the SRF gun. To determine the optimal voltage ratio between the cathode and focusing electrode, various DC gun field maps were generated and used as accelerating gradient sections in GPT simulations. The beamline layout, as shown in Figure 63, was incorporated into the

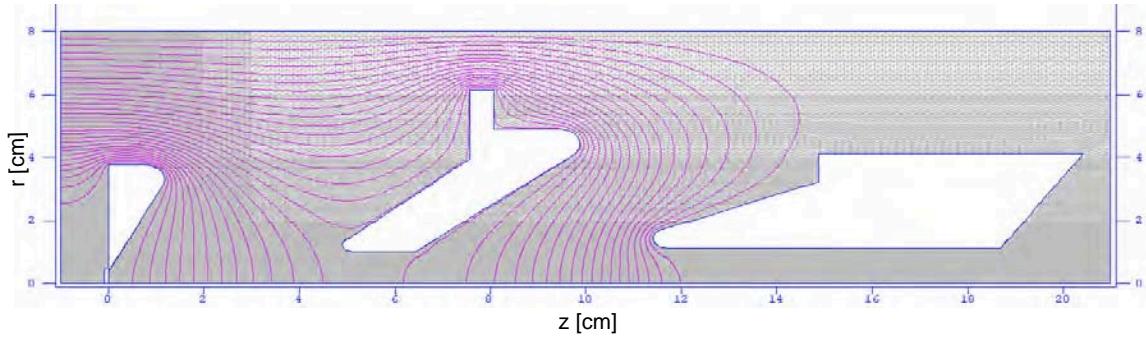


Figure 65. The Poisson model for the Stanford DC Gun. The left-most structure is the cathode and Pierce cone. The middle electrode is shaped to provide beam focusing as the beam is accelerated toward the grounded electrode (final structure on the right). The pink lines mark equipotential lines.

GPT model and various bunch charges propagated to determine the optimal gun setup for performing the desired measurements. The results of these simulation runs are shown in Figures 66 and 67.

The simulation runs show us important values as the beam evolves moving through the DC gun structure. In Figure 66, we see that for the five electrode configurations (where the first number indicates the percent of the total voltage drop between the cathode and focusing electrode and the second number is the percent voltage between the focusing electrode and the ground electrode) the beam gains energy as expected. In these runs the maximum voltage potential was established as -160 kV, commensurate with the conditioned status of the system at the time. These simulations also reflect relevant parameters expected in actual experiment—a 5 nC bunch charge liberated in a 4 ns pulse (due to pulse length of the Minilite II drive laser) from a 4 mm radius spot size at the cathode and a uniform beam distribution.

If we now focus our attention on Figure 67, we see that it is possible to transport the entire bunch charge as described through the gun structure. Of primary interest is the effect of space charge on the bunch. In the 25-75 case, the bunch drastically expands through the focusing electrode (low focusing, and still a lower

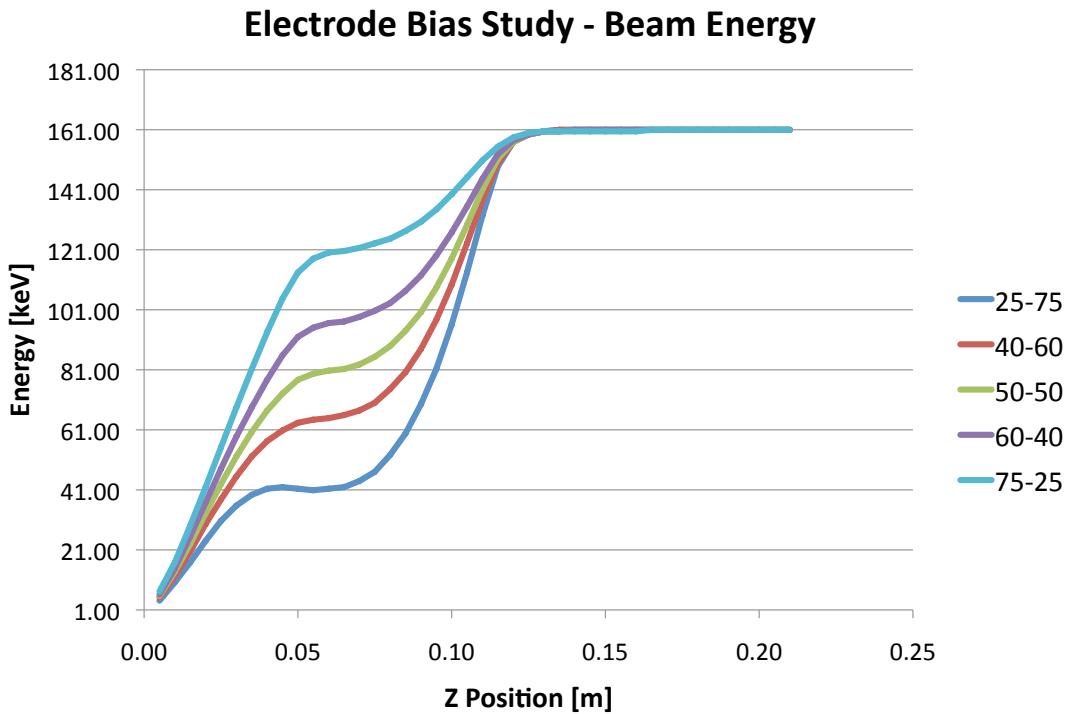


Figure 66. As the voltage ratio between the focusing electrode and cathode are changed, the beam falls through a different potential field. The plots show the effect of changing this ratio for a constant maximum voltage of -160 kV, illuminated laser spot, and bunch charge.

energy beam) and exits the gun at nearly the full aperture of the transport system. From this graph, we can see that it is desirable to have as much of the potential drop across the cathode-focusing electrode gap as possible. In realistic terms, the voltage divider is only capable of delivering approximately 57% of the potential between the cathode and focusing electrode, putting the expected beam envelope close to that of the 60-40 case. We find this setup acceptable as it transports a beam from the gun diverging at approximately the same rate as all cases considered with the exception of the 25-75 case. The location of the normal conducting solenoid, immediately after

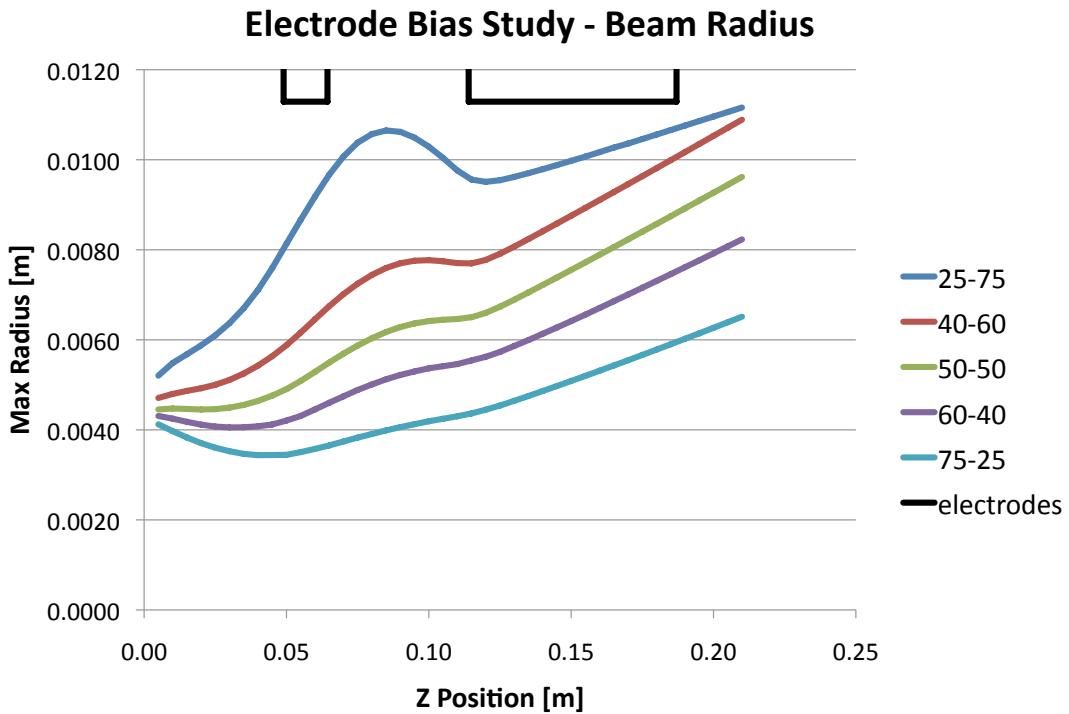


Figure 67. As the voltage ratio between the focusing electrode and cathode are changed, the beam envelope changes. The plots show the effect of changing this ratio for a constant maximum voltage of -160kV, illuminated laser spot, and bunch charge.

an isolating gate valve, is as close to gun exit as practicable, and we can expect to be able to focus the majority of the beam to the diagnostics further down the beam line.

Our first experiments focused on attempting to measure the cathode's quantum efficiency. Quantum efficiency is a ratio of the number of electrons generated per incident photon. To measure this quantity, we integrate the voltage signal from the Faraday cups that measure the current passing through a known resistance providing the bunch charge. Simultaneously, we measure the pulse energy of the laser pulse incident on the cathode surface and assume that the transmission properties of the transport line remain constant and that the full pulse energy reflected from the laser mirror is incident on the cathode surface. As an example using actual data (266

nm UV wavelength, pulse energy of $170 \mu\text{J}$, 50% transmission, $V_{FC1} = 1.68 \text{ V}$, $V_{FC2} = 7.3 \text{ V}$, $\sigma_{FC1} = 7.18 \text{ ns}$, $\sigma_{FC2} = 12 \text{ ns}$, and $\mathcal{R} = 10 \text{ k}\Omega$, and assuming the voltage pulses are triangular), where ‘‘FC’’ indicates the Faraday cup the measurement was taken from, we find the quantum efficiency of the system to be

$$\begin{aligned}
\frac{q}{e} &= QE \frac{\frac{E}{2} \lambda}{hc} \\
QE &= \frac{2qhc}{eE\lambda} \\
QE &= 0.5 \left(\frac{V_{FC1} \sigma_{FC1} + V_{FC2} \sigma_{FC2}}{\mathcal{R}} \right) \frac{2hc}{eE\lambda} \\
QE &= \left[\frac{(1.68 \text{ V}) (7.18 \times 10^{-9} \text{ s}) + (7.3 \text{ V}) (12 \times 10^{-9} \text{ s})}{10,000 \text{ k}\Omega} \right] \times \\
&\quad \left[\frac{1240 \text{ eV nm}}{(170 \times 10^{-6} \text{ J}) (1.602 \times 10^{-19} \text{ J/eV})^{-1} (1.602 \times 10^{-19} \text{ C}) (266 \text{ nm})} \right] \\
QE &= (9.966 \times 10^{-12} \text{ C}) (2.742 \times 10^4 \text{ C}^{-1}) \\
QE &= 2.732 \times 10^{-7} .
\end{aligned} \tag{IV.18}$$

As the first measurements were made, the quantum efficiencies and bunch charges were much lower than expected for a photothermal cathode [36], [37], [38]. Figures 68 and 69 show plots of quantum efficiency versus heater voltage and quantum efficiency decay over time for the Stanford DC gun.

To compile the data shown in Figure 68, the cathode was allowed to stabilize at a specified heater voltage for a minimum of 10 minutes prior to data collection. Beam transport was optimized (maximizing $V_{FC2} + V_{FC1}$) for a constant gun voltage of -145 kV, which was slightly below the maximum conditioned voltage at the time. Bunch charge and pulse energy were recorded at 30 second intervals for 3 minutes and averaged to compile a quantum efficiency for a particular heater voltage. Laser pulse energy was constant at approximately $170 \mu\text{J}$ throughout the experiment. The illuminated area on the cathode is estimated to be approximately 1 cm in diameter which would illuminate the entire active area of the cathode. In reviewing the data and comparing to estimates and measurements of quantum efficiency for comparable cathodes, our values were found to be at least two orders of magnitude below

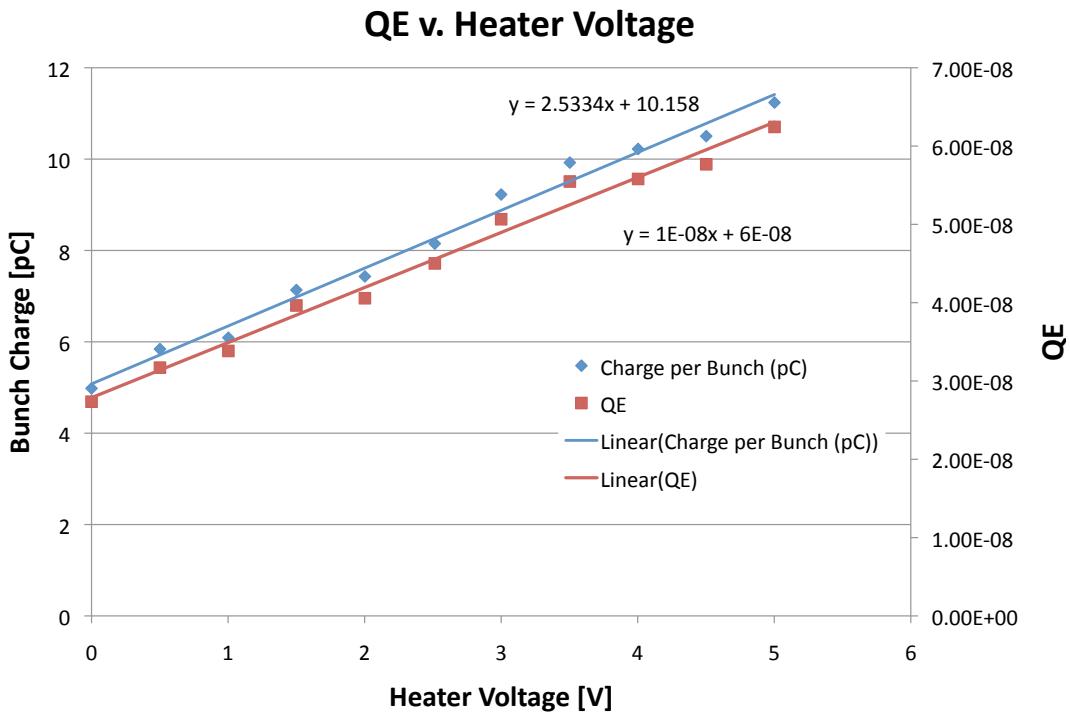


Figure 68. Quantum efficiency of the DC gun as heater voltage is increased.

expectations. Prior to performing further experiments, we focused on attempting to increase laser energy transport and further conditioning of the cathode and gun.

After application of the improvement tasks, the DC gun photothermal cathode finally began to approach quantum efficiencies measured in other systems. Bunch charges were measured commensurate with the improved quantum efficiency and higher pulse energy (now delivering approximately $250 \mu\text{J}$ to the cathode surface). Additionally, the $10 \text{ k}\Omega$ resistors previously used to measure voltage on the Faraday cups were replaced with 50Ω resistors for better signal matching to the oscilloscope. Figure 69 shows a quantum efficiency decay plot for the cathode after performing a brief surge to full heater voltage (5.0 V) for 30 minutes and then allowing the cathode to cool for 30 minutes prior to taking data. We see the expected exponential decay, but the cathode did not reach a minimum value as we expected. Our expectation,

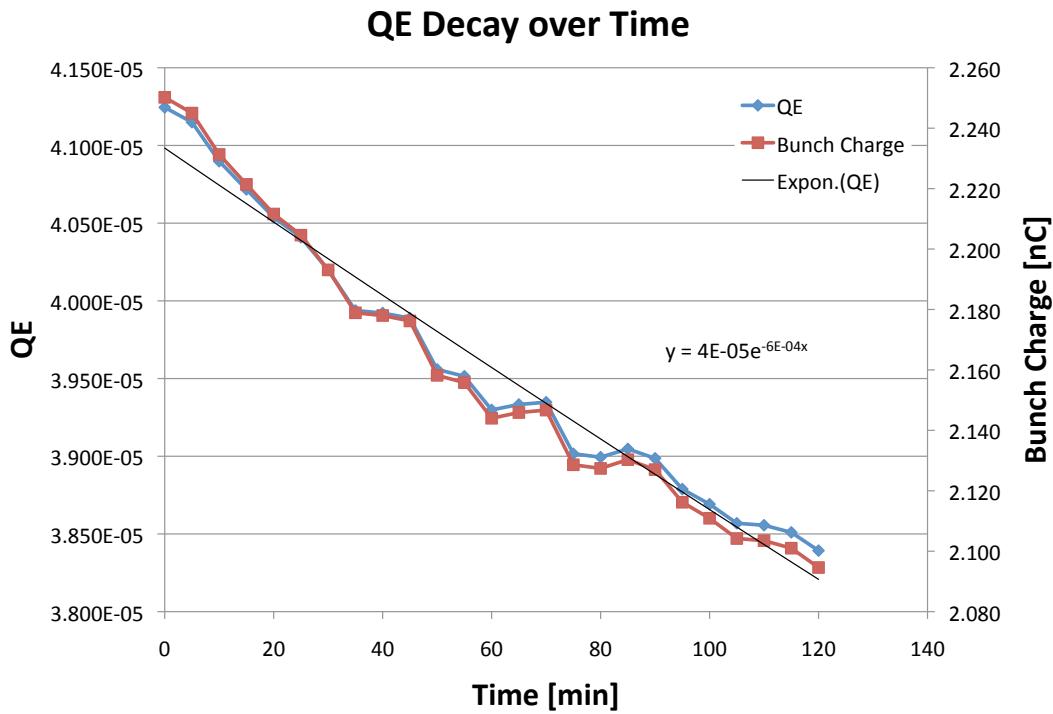


Figure 69. Quantum efficiency decay of the DC gun cathode held at constant voltage below the replenishment voltage.

given the age and history of the cathode, was that the decay plots should start to reach a minimum fairly quickly after the heater was turned off. It is possible that these type dispenser cathodes are much more robust than we anticipated. Unfortunately, due to external factors, we have not been able to revisit this to extend the measurement period in hopes of observing a minimum.

As experiments were performed using the DC gun, numerous issues were identified that may add uncertainty to the results obtained thus far. With the drive laser wavelength in the UV, we have been unable to determine where the electrons are being generated. When using the UV beam (266 nm), all known metals in the region of the cathode (cathode surface, grid, Pierce cone, all other electrodes) will photoemit. Removing the second doubler crystal allows for operating the laser in the green region (532 nm) which is below the work function of copper. Unfortunately,

we were unable to validate laser alignment at 532 nm due to laser crystal issues and could not rule out photoemission from the copper structures in the gun.

Transporting the electron beam a significant distance allows for beam expansion and possible losses in transport, especially at low energies. Most quantum efficiency studies are performed with a small cathode-anode gap and extremely short transport distances to ensure capturing all generated charge in the collection mechanism. With the DC gun, this is problematic due to the high voltages involved. However, in this device, it is possible to measure the replenishment current flowing into the cathode region. To improve our ability to measure where the electrons in the beam were being generated, the Eimac cathode's separate (and isolated) electrical connections, shown in Figure 70, have been instrumented with fast current transformers to capture the replenishment charge flowing into the various parts (Pierce electrode, cathode, and grid). The concept has been verified at low voltage without the drive laser, but we have not been able to test at design gradients due to an ongoing radiation stand down at our NPS facilities.

Since the Eimac cathode used in the DC gun system is the one that was shipped installed in the gun from Stanford and was exposed to air for more than two years before being placed under vacuum, it is possible that the cathode has been significantly contaminated. No attempt to reinitialize the cathode was made for the results presented here. A modified reinitialization was attempted later, but no significant improvement was seen. To better understand the cathode and its performance, and possible effects on an SRF cavity, separate, non-beam generating experiments were conducted [41]. To incorporate lessons learned and minimize nonessential variables, a lower voltage (100 kV) test cell with a simplified cathode-anode geometry is being manufactured that will incorporate the improved measuring scheme and the upgraded drive laser system for testing various photothermal beam generating schemes.

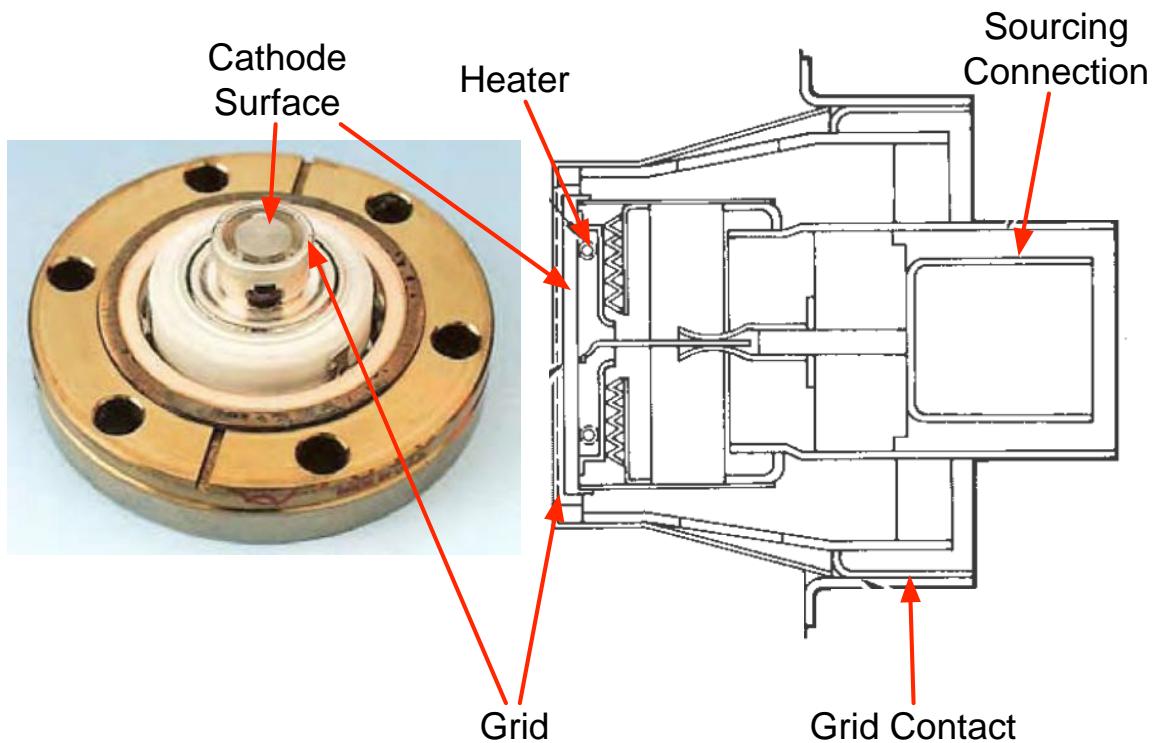


Figure 70. The general layout of an Eimac electron source is shown. The picture on the left is an actual device, removed from its vacuum system. Connection to the lab system is by conflat flange with copper seal. The related diagram on the right shows the internal wiring and isolation of the cathode sourcing, grid bias, and heater systems. Not shown is the Pierce electrode, which is part of the gun internal structure—it would be located very near the grid but still electrically isolated. After [39], [40].

5. Booster Configuration

The first injector system modeled was the booster configuration. In this configuration, it was intended that the source beam would be generated using the former Stanford DC gun at an energy of approximately 150-200 keV. The booster would provide an additional 1.2 MeV for merging into the primary accelerator. Of primary difference with the previous DC gun simulations, the beam pulse length is greatly decreased. In the previous case, a 4 ns pulse was used, but this corresponds to two complete RF cycles of the 500 MHz booster and is not ideal. To correspond to commercially available UV drive laser specifications, a “long” drive laser pulse of 40 ps is used with a reduced bunch charge (1 nC). There are drive lasers available with shorter pulses and higher repetition rates, however this pulse length corresponds to the drive laser being used in the current 500 MHz gun experiments.

Since the booster is of primary concern for this study, we assume the input beam is uniform, monoenergetic, and has an initial emittance of zero to remove any effects that might be present from the electron source. The beam is initiated at the closest point of connection from the booster to the gun section. Space charge forces are not applied until after the beam reaches this earliest connection point (62.7 cm before the booster cavity nose cone). The simulated beam line is terminated at a point approximately 3.0 m from the booster nose cone to compare results from both Parmela and GPT, and to correspond to the beam line as originally proposed. These simulations also serve to compare the two code results on similar beam lines.

a. *Acceptance*

Acceptance as defined in the booster configuration under consideration will be defined as the cavity phases for which an injected beam can successfully be propagated through the cavity. Obviously, there are injection techniques that can be used to adjust the cavity acceptance, however these methods can only change the acceptance by a little compared to the overall acceptance of the cavity. To simulate the cavity acceptance, multiple runs, using the same beam parameters and only varying

the cavity phase, we can obtain general plots for the optimal phasing of the booster cavity for charge transport, beam size, emittance, etc. Figure 71 shows plots of the cavity acceptance for a 200 kV input beam as described previously.

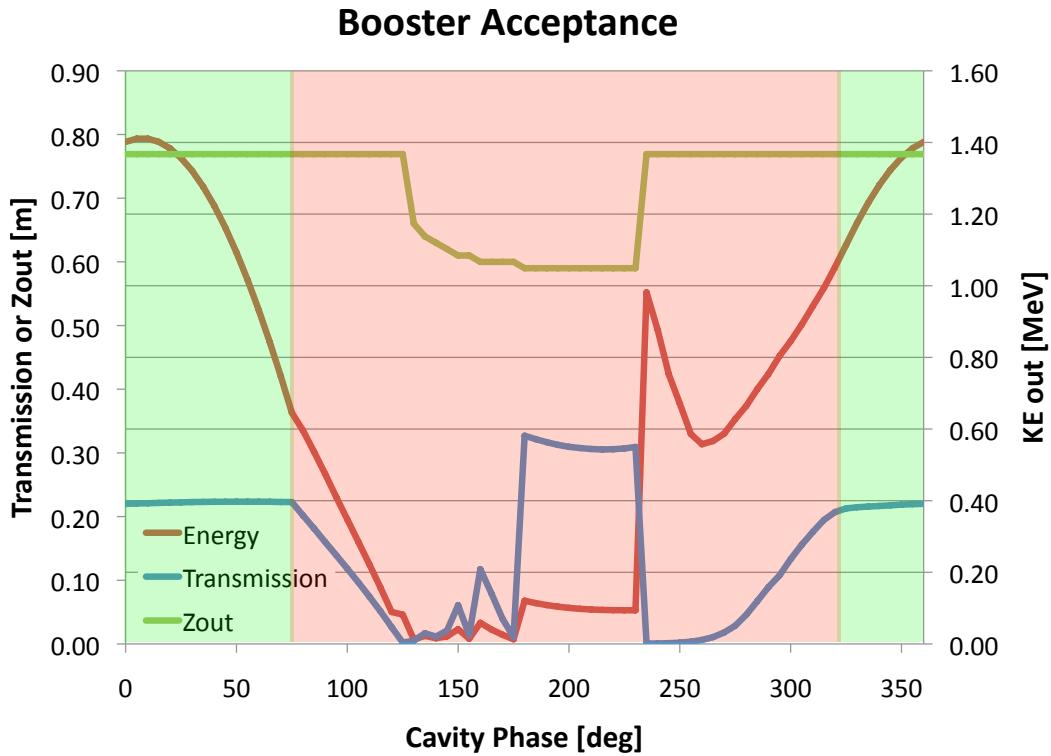


Figure 71. Acceptance of the NPS 500 MHz booster cavity for a 200 kV input beam. The green shaded area indicates “acceptable” beam performance characteristics while red indicates beam that has transport problems through the cavity.

In Figure 71, the “acceptable” cavity phases have been highlighted in green. To come to this conclusion about “acceptability,” there are two main data points that are considered. First is whether the beam is transported to the expected z position, in this case the center of the superconducting solenoid just after the booster cavity at a z position of about 0.79 m. Additionally, there should be a significant population of particles that are transported to the desired z location. The simulation software will calculate parameters for a single particle if it can successfully transport

only one particle through the desired geometry; however, these type results are not useful. From Figure 71, we see that the cavity, at best, transports about 22% of the initial population (of 30,000 particles) to the desired z position. Within the phase range of $-40^\circ \leq \phi \leq 75^\circ$, essentially the entire bunch that makes it through the drift to the cavity and through the iris of the nose cone and can be propagated to the first focusing element of the beamline. Other than the energy plot plotted in Figure 71 that reflects the phase of the cavity, we do not yet have any information about the beam itself.

b. Beam Transport

To obtain data about the beam itself, it is necessary to transport the beam through the desired beamline and observe its performance characteristics. In our case, we are primarily interested in the beam energy, radius, and emittance. Once the beam has left the accelerating structures, its average energy does not change unless beam scraping occurs. The internal beam energy will change as space charge accelerates the head and decelerates the tail, resulting in energy spread within the beam. With respect to the beam's radius, there are two primary measurements of interest. The first is the average beam radius which represents the primary beam that is transported. The second is the maximum radius which takes into account the particles that have been ejected from the core beam and make up the beam halo. While the beam size can be adjusted after the primary accelerator fairly easily, emittance after the gun/injector, in general, cannot. Unfortunately, without discarding portions of the beam, once emittance growth has occurred, it usually cannot be reduced again. It, therefore, behooves us to choose a launch phase where the beam does not have excessive emittance. Figure 72 demonstrates the effect of cavity phase on both beam radius and emittance at the booster exit. Recalling that the peak of the energy plot is at 0° , the ideal phase for radius and emittance will probably be less than 20° .

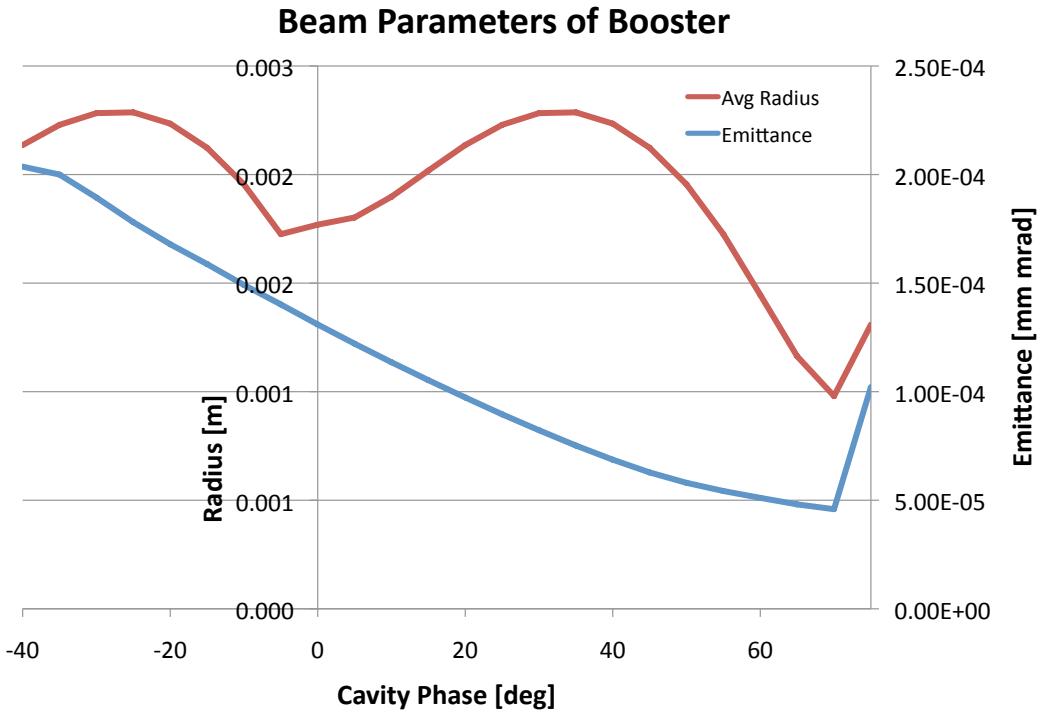


Figure 72. Emittance and beam radius dependence on cavity phase of the NPS booster cavity. The sharp increase at the far right is due to poor transport at that phase value.

Figure 73 shows the 200 kV beam evolution as it is propagated through our notional beam line from initiation, through the booster cavity, the superconducting solenoid, the normal conducting solenoid to a point approximating where the merger to an accelerator structure or other diagnostic would be placed (approximately 4.8 m from the beam launch). Transport parameters (solenoid strengths) were determined using the methodology discussed earlier, attempting to preserve as much of the beam as possible. After optimization through the beam line, the parameters of interest were found to be: $\phi = 29.41^\circ$, $B_{SC-pk} = 280$ mT, $B_{NC-pk} = 28.7$ mT, where “SC” indicates the superconducting solenoid and “NC” indicates the normal conducting solenoid. The transport envelope using these optimized values is shown in Figure 73. We see that the beam radius increases significantly from initiation to

the booster nose cone iris, where the beam loses more than 1/3 of the charge. From this point, we see the remaining beam is fairly successfully transported to a beam focus at the target. The divergence of the maximum radius from the average radius is an indication that the combination of the booster cavity and superconducting solenoid generate significant halo that impact the beam pipe just before the normal conducting solenoid. Propagation from the normal conducting solenoid to the target is uneventful other than generation of halo from the back of the beam where it is overfocused by the normal conducting solenoid at about the $z = 4.0$ m point.

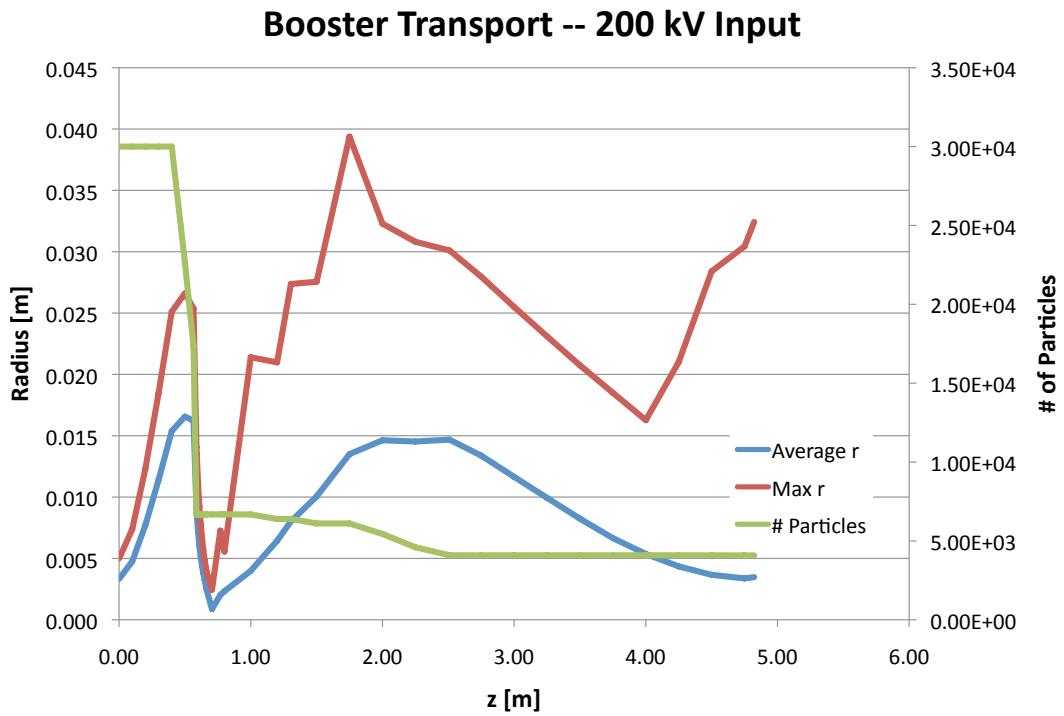


Figure 73. Beam transport envelope for the NPS booster cavity with a 200 kV input beam.

In an attempt to increase beam transport and to counter the effects of space charge in the initial drift prior to the booster cavity, a second normal conducting solenoid with the same field map as the beam line version is placed between 20 and 30 cm from the gun connection (which is also 20-30 cm before the booster cavity

nose cone). We optimize for the solenoid positioning and peak field strength using the transport parameters (ϕ , B_{SC-pk} , and B_{NC-pk}) from before for comparison. The new transport envelopes are shown in Figure 74 and show significant improvement in booster cavity performance. Optimization places the center of the input solenoid at 22.7 cm from the beam start with a peak field strength of 66.9 mT. This position and peak field result in a maximum beam radius of about 1.3 cm as it passes through the booster cavity nose cone. From the radius plots and loss of particles shown in the particle plot, it now appears that the superconducting solenoid is overfocusing the beam and that further optimization would result in nearly full transport of the beam. Regardless, the addition of this one element increased beam transport by a factor of more than four.

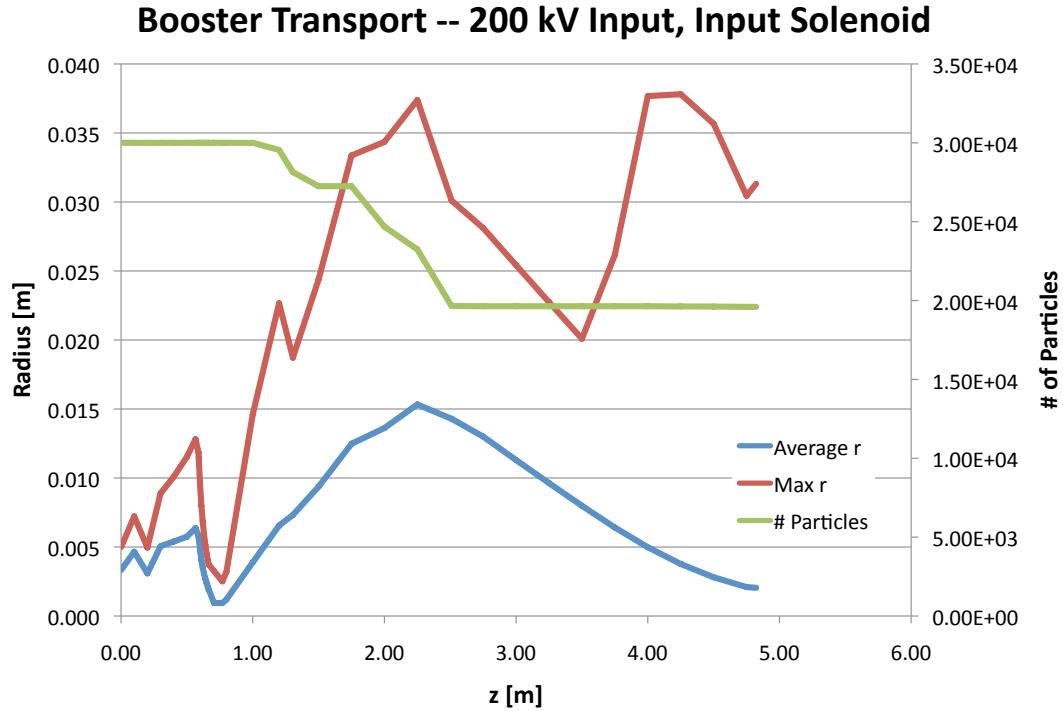


Figure 74. Beam transport envelope for the NPS booster cavity with a 200 kV input beam and the addition of a normal conducting solenoid prior to booster injection.

The impact of adding this solenoid is important to the transverse emittance. Figure 75 shows the two optimized beams overlayed for comparison; the only difference is the addition of the pre-booster solenoid. The blue plot represents the addition of the starting solenoid, which the red plot represents only a drift until the booster cavity. From these plots, we see the solenoid initially increases emittance, but the emittance growth seen in the red plot going through the booster cavity is mitigated by the effect of the starting solenoid. This emittance improvement (almost a factor of 2 over the no-focus injected beam) is maintained almost through the rest of the entire transport line. The non-compensated beam having an emittance of $66.3 \mu\text{m}$ compared to the solenoid compensated beam emittance of $40.8 \mu\text{m}$. Neither emittance is particularly noteworthy, but considering the less than ideal initiation, the fact that it is not larger is of note.

From this study, it appears that the NPS cavity is suitable as a booster, but either the input drift should be minimized or some emittance compensation device (such as a solenoid) should be included in the design prior to the booster. However, an injection energy of only 1.4 MeV may present a challenging merge problem for the main accelerator. This issue is not taken up in this study, but left for future work. We now depart the booster concept briefly to examine the use of the NPS 500 MHz cavity in the gun configuration.

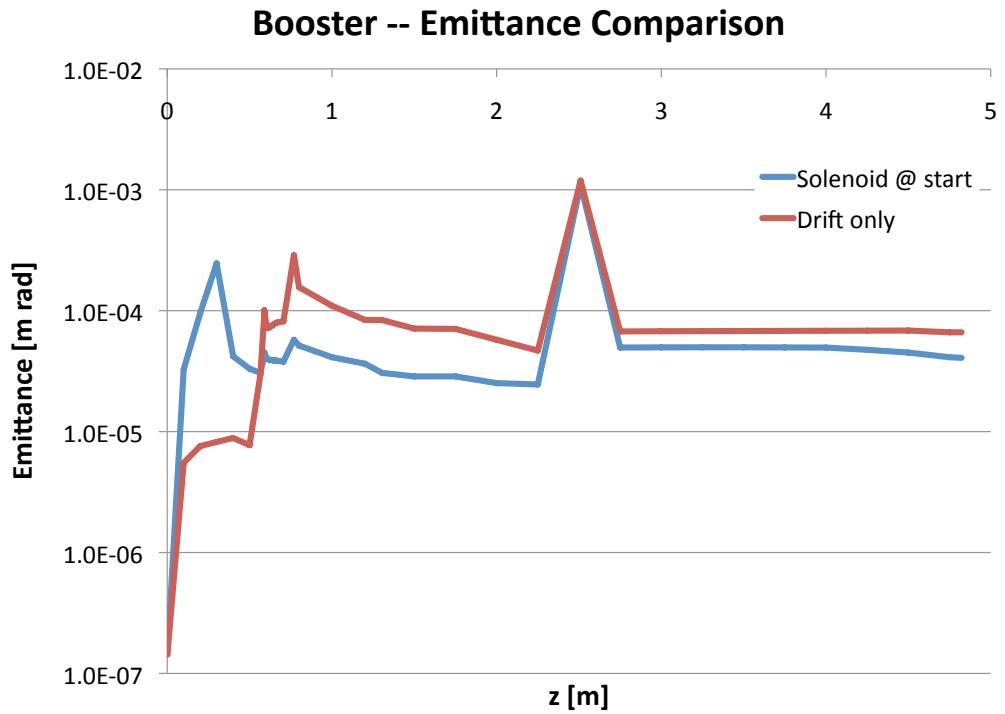


Figure 75. The addition of a solenoid prior to the booster results in higher emittance at the gun, but lower emittance at the end of the transport line. The final emittance value for the drift only simulation is a significant underestimate as the gun nose cone has filtered out the high emittance particles and the transported final bunch charge is significantly lower than with the addition of the input solenoid.

6. Gun Configuration

Insertion of a cathode stalk into the booster allows the same cavity to be used as a gun. The type of electron source is dependent upon the type of cathode installed. The NPS SRF gun has been designed with multiple cathode types in mind, including field emitter arrays and photoemitter cathodes (metal and semiconducting). Turn-on fields for gated diamond field emitter arrays have been described as anywhere from 1 MV/m [23] to 15 MV/m [7]. As these are fairly new devices that have typically only been operated in test stands with small voltages and small cathode-anode gaps, their performance in realistic gun environments is unknown. If we assume the turn-on voltage in a real application will be bounded by these values, and assume 15 MV/m is necessary for full current development from a field emitter tip, then we see from Figure 17 that fields are acceptable for field emitter array cathode use for all retraction positions (relative to the nose cone) forward of 6.5 mm. The 6.5 mm retraction position represents the maximal retraction for ensuring turn-on across a 1 cm diameter array cathode as electric field values drop significantly as they are measured radially from the electric axis. For photoelectric cathodes, a drive laser is required along with the laser port into the beam line and a mirror to direct the laser beam to the cathode.

For both simulation packages, the beam line was modeled (elements placed) based on the system installed at Niowave. Figure 76 shows the layout and provides the as-built measurements. The only difference in the model input files is that the distance from the nose cone to the cathode surface has to be accounted for as all measurements in the beam line use the nose cone as the reference point.

a. Acceptance

Using a standard 1 nC, uniform electron beam, we first determine the expected acceptance of the gun for various cathode positions. Figures 77 and 78 show these plots and pertinent values as a function of the launch phase of the electron beam. These figures are fairly dense with information, so we take some time to analyze them.

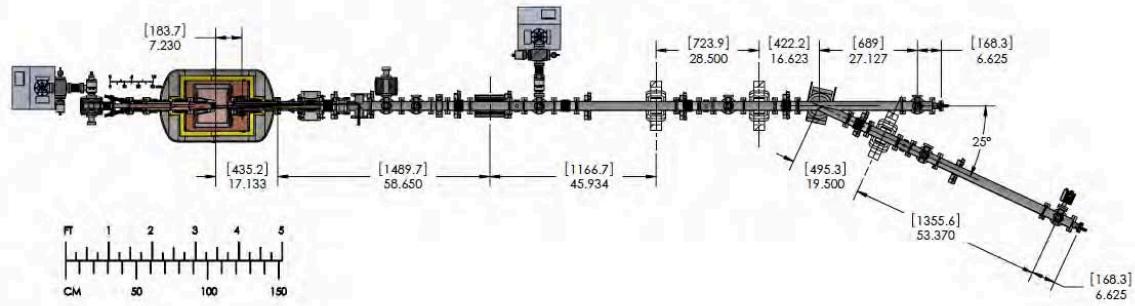


Figure 76. The NPS 500 MHz gun diagnostic line has most of the instrumentation necessary to fully characterize the expected beams from the gun. Numerous steering coils, diagnostic pots (each of which includes a Faraday cup, an yttrium aluminum garnet crystal for beam imaging, and space for future pepper pots for emittance measurements) and quadrupoles are installed for ensuring proper transport. A dipole magnet and spectrometer leg are included for beam energy and energy spread measurements as well.

Looking at the top of Figure 77, we see that all of the cathode positions have about the same phase acceptance window ($85^\circ \leq \phi \leq 225^\circ$) for the cavity voltage applied (1.24 MV). We also see from this plot that the focusing at the cathode provided by retracting the cathode behind the nose cone is critical—until the cathode is retracted about 5 mm, full transport of the bunch charge is impossible. Once the cathode is retracted approximately 6.5 mm, full transport of a 1 nC bunch becomes possible at nearly all phase angles. Moving to the energy plots in the bottom of Figure 77, we see that as the cathode is retracted, the peak of the energy plots are shifting slightly earlier. This can be accounted for by the additional time of flight required by the bunch to reach the high field regions in the gap. Peak average beam energy is affected by the cathode position, but only varies by about 10 keV over the 9 mm of cathode travel. The spikes in two of the plots near the turn off phase are artifacts of how GPT counts particles. Since they are far beyond the peak energy where we plan to operate the cavity, they can be ignored. The radius plots in the top of Figure 78 provide correlation with the transport plots above them. The ragged lines

are indicative of particles being lost during transport. As the average radius plots remain fairly smooth for retractions positions beyond 5 mm, it can be assumed that the losses in these beams are small compared to the entire bunch charge. The spike in radius seen in the 9 mm retraction position in this subplot can be attributed to being beyond the acceptance cutoff of the cavity and ignored. In the bottom plot of Figure 78, we see that the launch phase chosen has an impact on the transverse emittance expected out of the gun as well. We also see that it is desirable to operate the gun in the region where the fields are building as this results in generally lower emittance values. From this plot, we further see that the cathode position again plays a role in emittance determination. For the 0 and 2.5 mm retraction plots, the ragged lines indicate again that particles are being lost—that cavity transport is acting as an emittance filter removing particles contributing to higher emittance as they strike the beam tube or fail to exit the cavity. If these plots followed the general trend of the other retraction values prior to transport losses, it is likely they would result in higher emittance values.

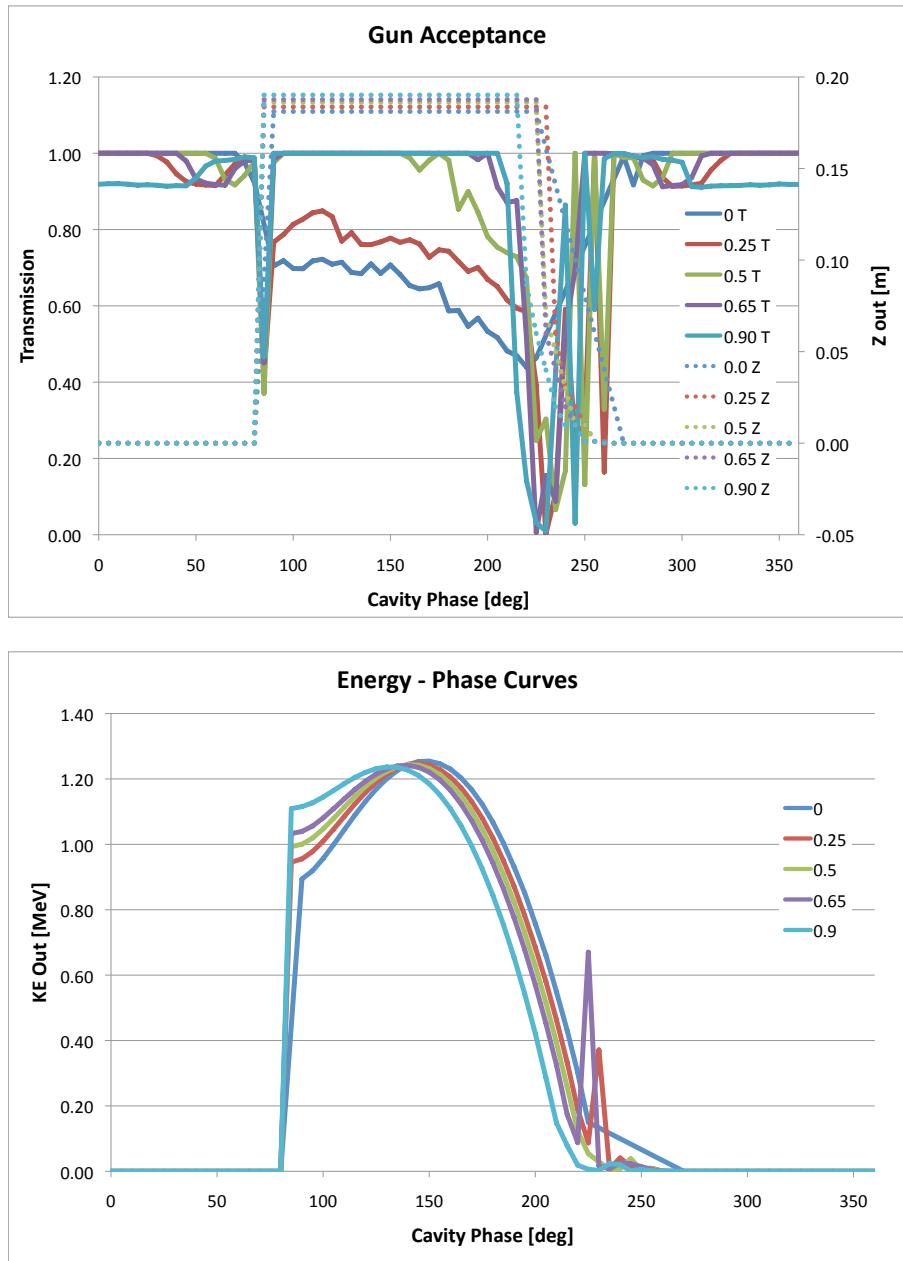


Figure 77. Acceptance plots of the NPS gun are plotted for five cathode positions. The positions are labeled with respect to their retraction behind the nose cone in mm (i.e., “0.0” is the cathode face flush with the nose cone and “0.9” is 9 mm retracted behind the nose cone). The top plot is the general transmission of charge through the cavity with the z position of the final screen passed included for clarification. The bottom plot is the output final energy versus launch phase plots.

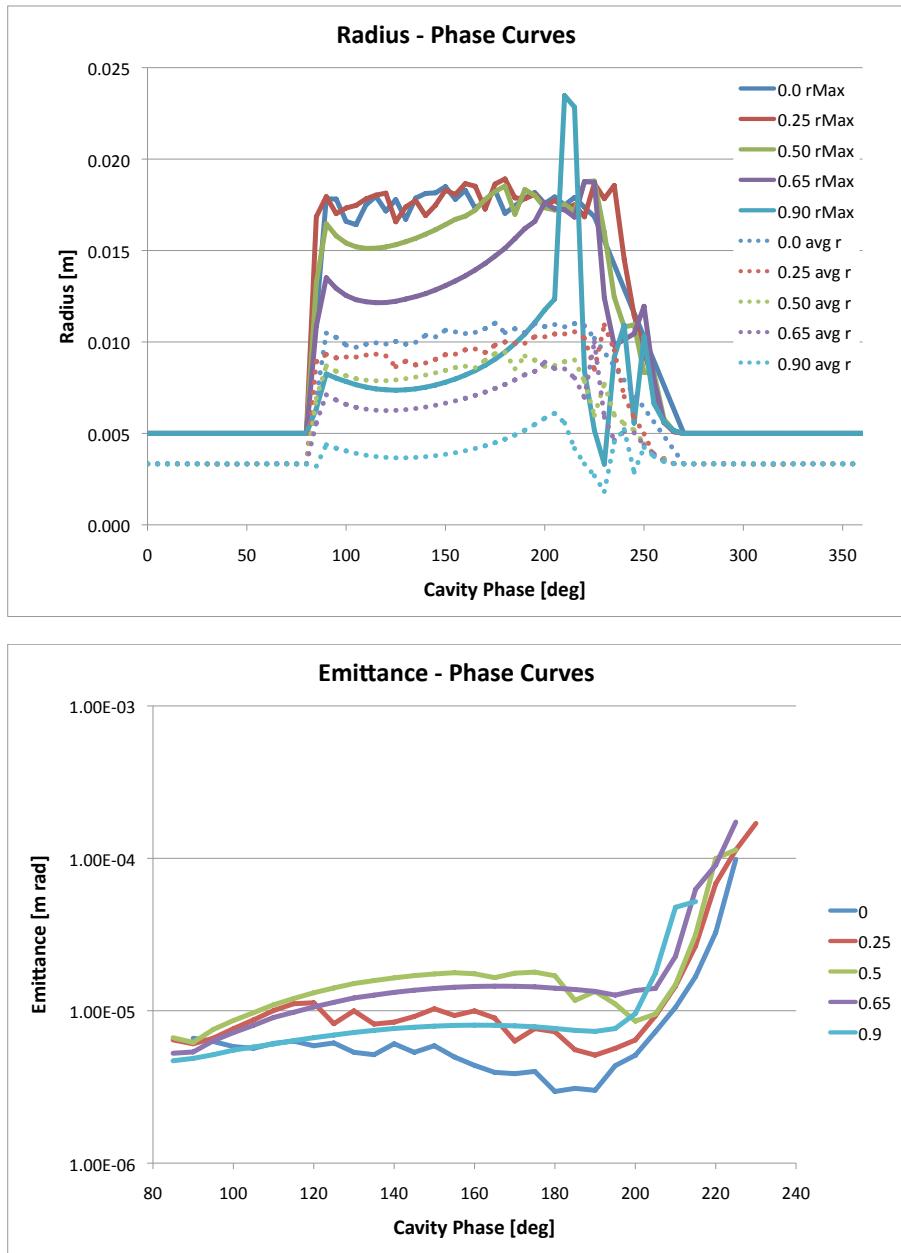


Figure 78. Radius and emittance dependence of the NPS gun on phase plotted for five cathode positions. The positions are labeled with respect to their retraction behind the nose cone in mm (i.e., “0.0” is the cathode face flush with the nose cone and “0.9” is 9 mm retracted behind the nose cone). The top plot is maximum and average radius at the final output screen as a function of phase. The bottom plot is transverse emittance as a function of launch phase.

To examine the impact of the pulse shape on the gun performance, we choose a single cathode position for analysis. The 6.5 mm retraction position was chosen as we know from Superfish studies that the entire cathode face remains above the expected turn-on voltage for field emitter cathodes, broadening the overall applicability of the beam studies performed. Using these more advanced beam shapes, we plot in Figures 79 and 80 the acceptance plots for a single cathode position to examine the effects of more realistic beams on the gun acceptance.

From Figure 79, we can see that the acceptance window for each of the beam types is essentially the same. Recalling that for the window to be valid, both the z_{out} position should be greater than 18 cm and the transmission should be close to 1.0. We find the cavity phase window for each beam type to be about $85^\circ \leq \phi \leq 225^\circ$. There is little difference in the energy plots. The difference in the rising plots is merely a sign of the $\delta\phi$ used in generating the plot (5°) indicating the actual leading edge for transport through the cavity is probably near 83° and the difference between the beam types is about 1° . Looking to the radius plots in Figure 80, we see that all three beams show similar behavior. The average beam size for the Gaussian and asymmetric beams appears less than the uniform beam due to the Gaussian shape radially—there is more charge in the core of the Gaussian and asymmetric beams at the outset. The overall beam sizes, using the maximum radius all follow very similar plots. Likewise, there is little difference in the beam emittances, shown in Figure 80, other than an apparent slight advantage for the uniform beam for earlier phase delays, but it only a difference on the order of $1 \mu\text{m}$. From these plots, it appears that the NPS cavity is fairly tolerant of the beam shape and no major impacts differing from the simplest uniform beam occur when more complicated beams are considered.

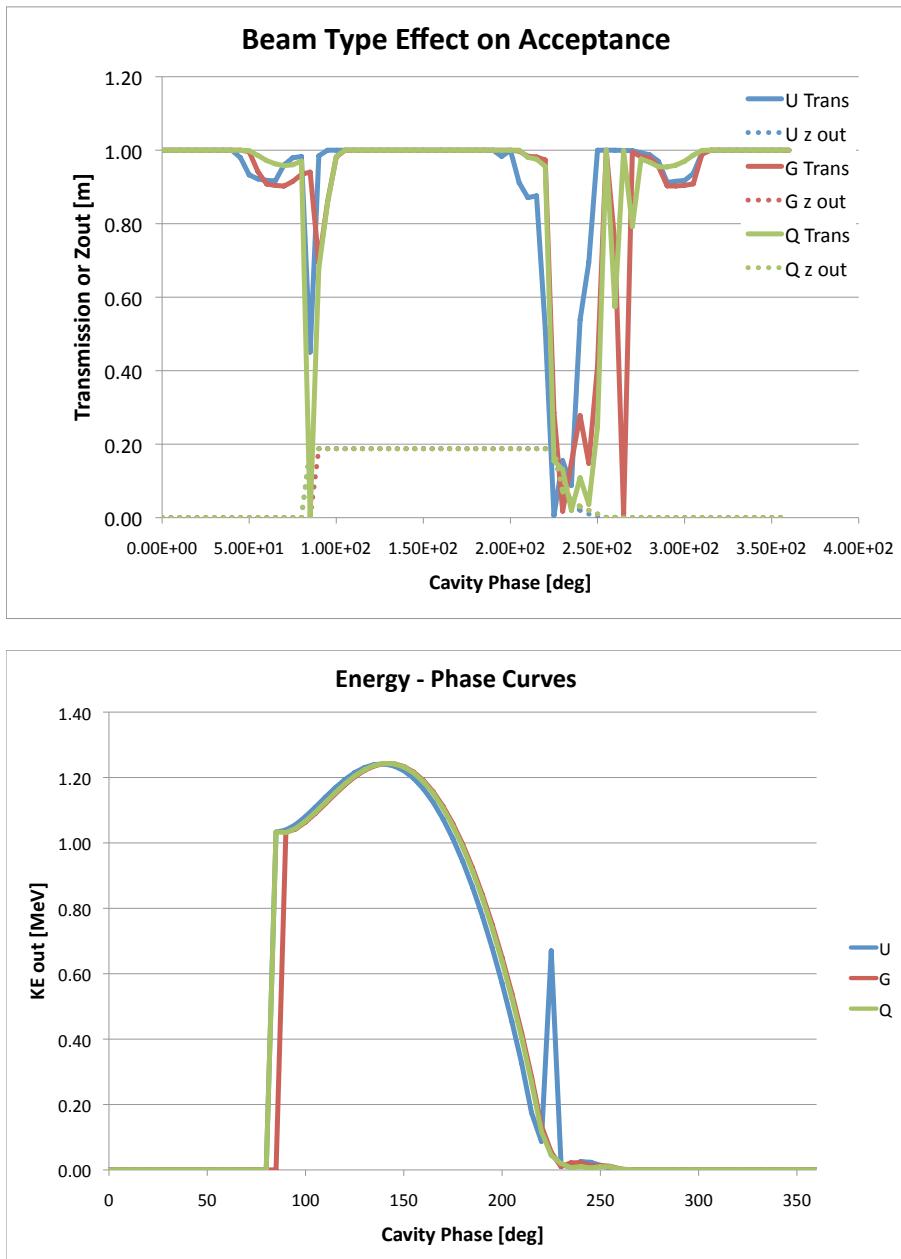


Figure 79. Various parameters of the NPS gun are plotted for three different beam types: Uniform, Gaussian, and asymmetric. The top plot is the general transmission of charge through the cavity with the z position of the final screen passed included for clarification. The bottom plot is the output final energy versus launch phase plots.

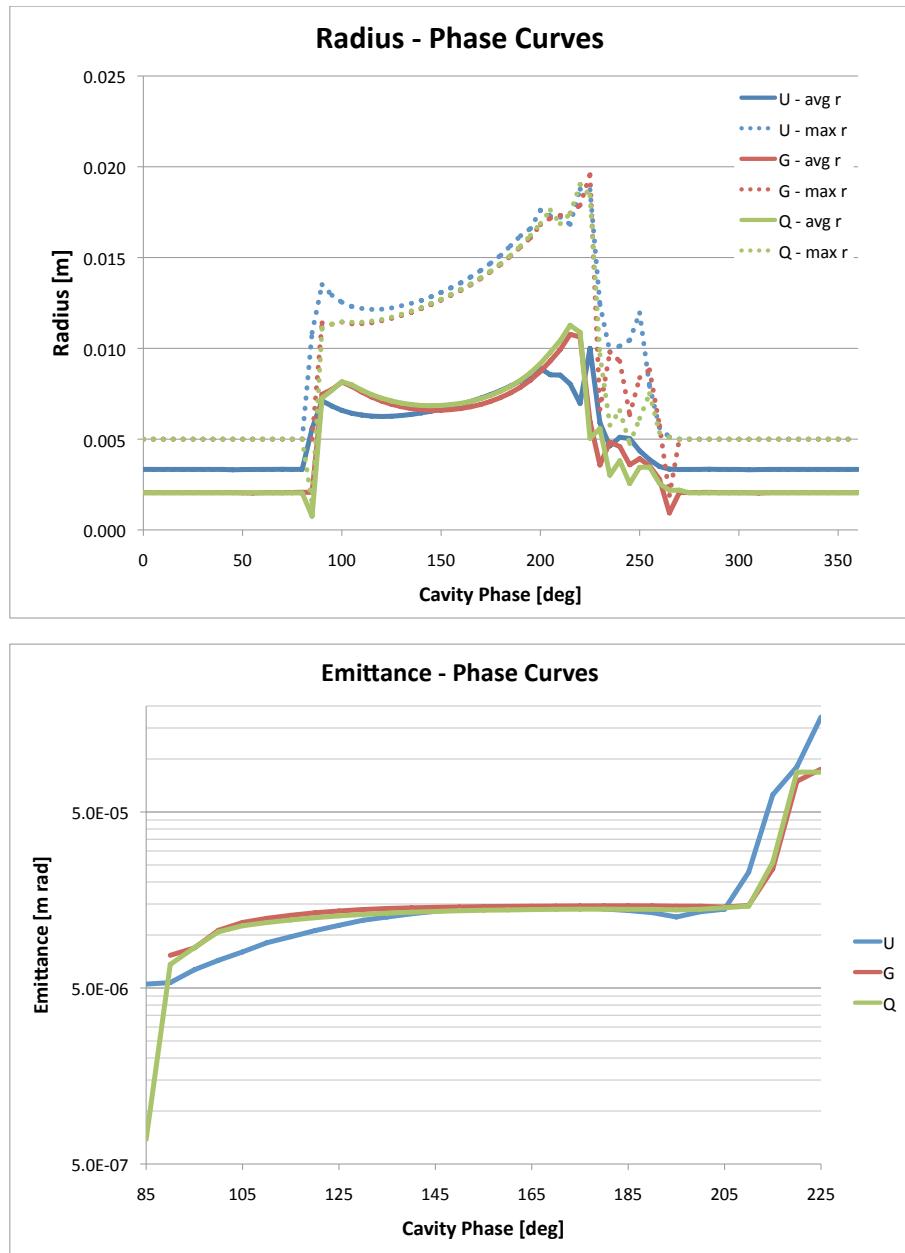


Figure 80. Radius and emittance dependence of the NPS gun on phase plotted for three different beam types: Uniform, Gaussian, and asymmetric. The top plot is maximum and average radius at the final output screen as a function of phase. The bottom plot is transverse emittance as a function of launch phase.

b. Beam Parameters

Now that the gun has been studied, further characterization requires propagation of representative beams to determine the beam characteristics that can be expected from the gun. Two primary optimization regimes are considered—optimizing for the smallest spot size at a given point (assumed to be a dipole or merger entrance 3.66 m from the nose cone in subsequent simulations) or for the minimum emittance at the same location. From the plots shown in Figures 77 and 79, we can expect the optimal launch phase to be near 150° to maximize energy for a cathode retracted 6.5 mm for both the emittance and spot size cases. Proceeding with the optimization scheme described previously, we see the final optimized beam envelopes in Figure 81 and the emittance evolution in Figure 82.

Looking at the plots in Figure 81, there is surprising similarity between the envelopes regardless of the optimization goal. The spot size optimization plots appear to delay halo formation in all the beam shapes until near the simulation termination except for the asymmetric case. In the emittance optimization, halo generation takes place for the uniform beam just before termination and the asymmetric beam about 60 cm before termination. From the beam visualization software, it can be seen that the halo formation is being generated by the lower energy tail of the beam being over focused due by the normal conducting solenoid. Of the three beam types, only the asymmetric beam appeared smaller (using the maximum radius as the measure of interest) when using the emittance optimization routine, but the terminal primary beam size was still smaller in the spot size optimization.

Beam emittance is a critical parameter for beam transport. In general, the smaller the number can be made, the easier transport becomes. Again, there was little difference between the two optimization routines when comparing beam to beam. The non-uniform beams were clearly better beams for transport arriving at the simulation termination with factors of 3 better (for the Gaussian beam) and 4 better (for the asymmetric case). From a prediction stand point this bodes well for using

the uniform beam as the primary simulation shape as it will tend to overestimate the beam emittance. Since the simulations do not take into account surface imperfections of the cathode, differing quantum efficiencies along the cathode surface, and were all started as cold beams (no initial energy or emittance), the uniform beam is probably the best the experimentalist can hope for in terms of predicted performance.

By the time the beams reached simulation termination, they had experienced significant longitudinal growth in all cases. For injection into another accelerating structure, we would like to have the beam fairly compact in order to ensure the beam does not develop any unwanted energy spread by sampling the accelerating fields too far out of phase from desired. To achieve this, we can shorten the transport line, however the line considered is probably close to a minimal design already. A buncher cavity could be installed to compress the bunch by passing the beam through an RF cavity near an RF zero crossing so that the head and tail of the beam receive kicks that would tend to compress the beam longitudinally. This was how the beam was compressed in the original Stanford FEL injector, however this scheme requires additional drift space for the compression to take place prior to injection into the main linear accelerator, expanding the total system length. The more compact solution is to add a booster to the gun and accelerate the bunch to a higher energy where space charge effects affecting the size of the bunch can be minimized.

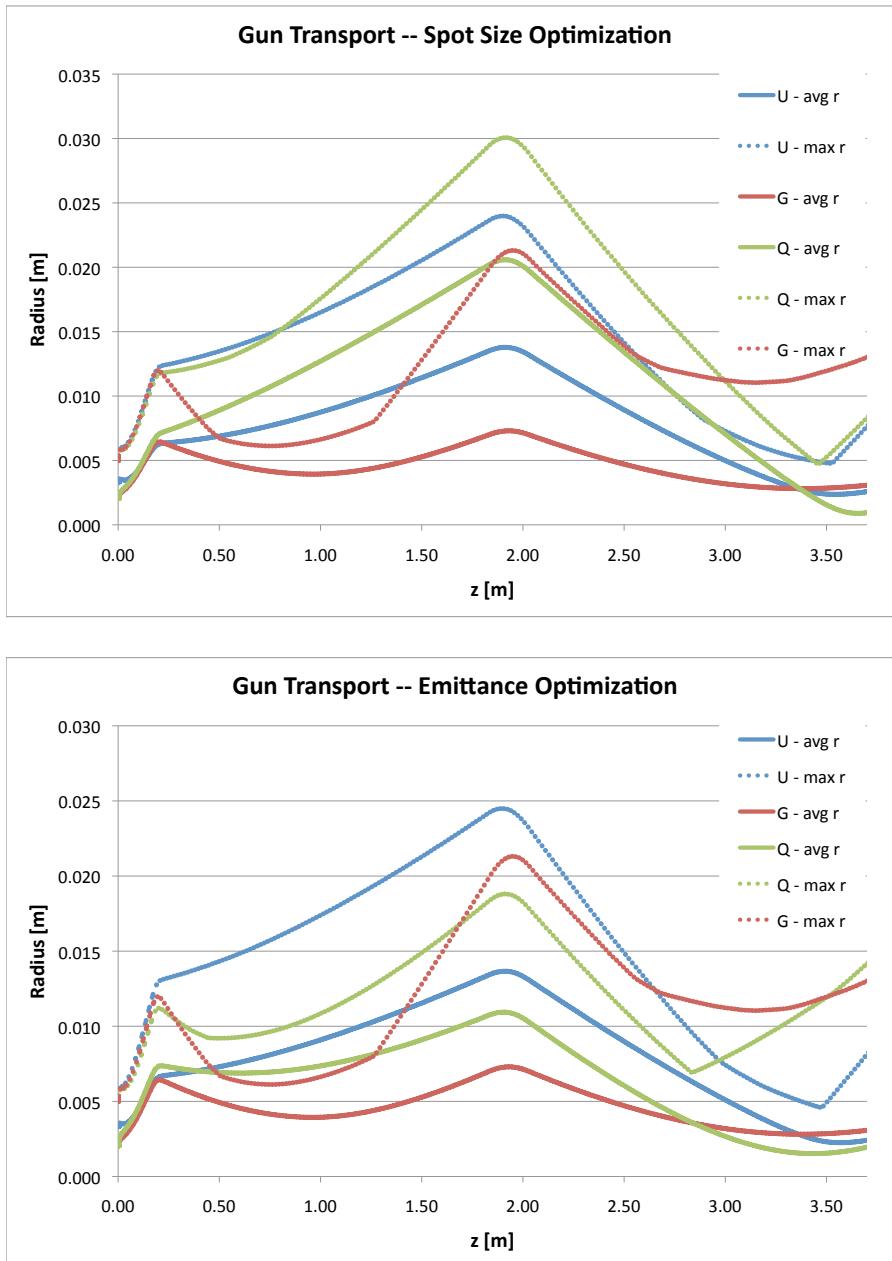


Figure 81. Optimized beam transport envelopes for uniform, Gaussian, and asymmetric beams initiated from the NPS gun. The top plot is based on optimizing for minimum spot size. The bottom plot is based on optimization for minimum emittance.

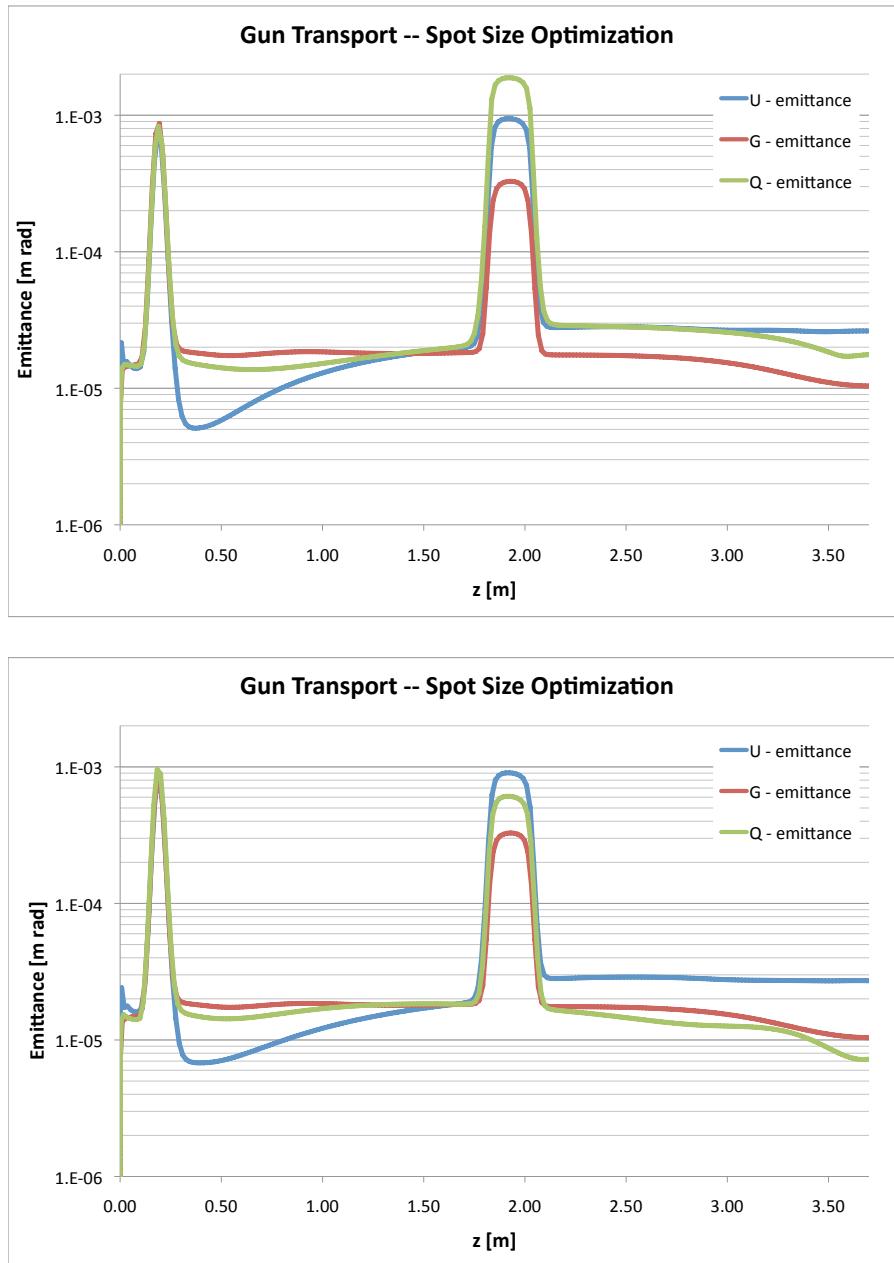


Figure 82. Optimized beam emittance evolution plots for uniform, Gaussian, and asymmetric beams initiated from the NPS gun. The top plot is based on optimizing for minimum spot size. The bottom plot is based on optimization for minimum emittance.

7. Injector Proposal

During the exercise of predicting the performance of the NPS 500 MHz gun, it became apparent that with this gun design, it was of utmost importance to increase the beam energy as quickly as possible in order to minimize the space charge effects experienced by the electron beam. There are many booster cavity designs already developed. However, considering that all fixtures and designs for manufacturing the quarter wave structure have already been developed, the author proposes and analyzes a similar end-to-end structure be used as a booster. Figure 83 provides a Superfish model of such a cavity.

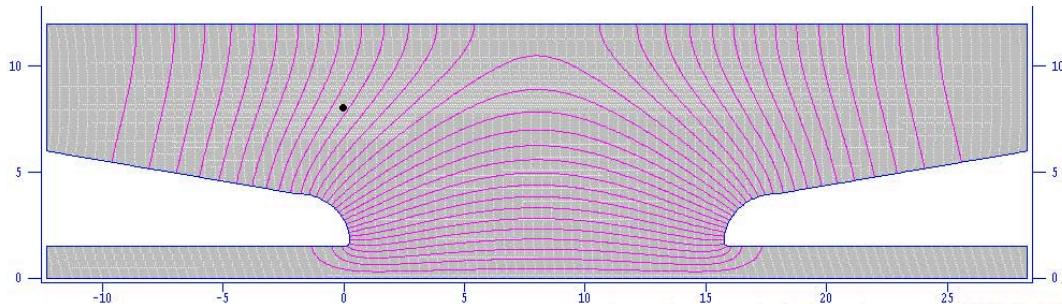


Figure 83. The first iteration of the double bell quarter wave booster cavity. This is a reflection of the NPS 500 MHz quarter wave gun, with simplified nose cone structures, taken along the end plane of the cavity. The fundamental frequency is 500.30 MHz and the pink lines follow electric field lines.

a. Cavity Development

To develop this booster structure based upon the NPS gun/booster geometry, there are two primary design goals: accelerate the beam generated from the NPS 500 MHz gun to an energy > 3 MeV for merge properties with the main accelerator and do so in as compact a package as possible at the same fundamental frequency. To accomplish these goals, some assumptions were made to aid the analysis. The first assumption was that the coaxial coupler currently installed in the NPS gun would be replaced with a coaxial coupler perpendicular to the beam line as recommended in the previous chapter. This provides RF power to the NPS gun. Second, the entire

assembly of gun and booster would be immersed in a combined cryomodule with the superconducting solenoids at the exit of the gun and the booster immersed in liquid helium to ensure they are below T_c . Third, to minimize additional drifts, RF power is assumed to be applied to the booster cavity via intracavity antennae. No modifications to the actual NPS gun geometry will be modeled; only basic modifications to the booster cavity were modeled.

As a first attempt at generating a booster cavity, the NPS gun structure (minus the cathode assembly) was simply reflected about the anode side of the cavity, as shown in Figure 83. The booster cavity was then modeled in Superfish, using the same maximum electric and magnetic field parameters as were used in designing the NPS gun structure. This reflection approach worked surprisingly well, resulting in a booster cavity with a fundamental frequency of 500.30 MHz. After adding the correct internal nose cone structure, the Superfish model showed negligible change in the fundamental frequency. The resulting accelerating fields and radial fields on axis and 1 cm off axis are plotted in Figure 84. Peak on-axis accelerating fields, E_z , are found near the two nose cones and peak at 21.0 MV/m. The large dip in on-axis accelerating field in the center of the cavity was concerning as it is a drop of 48% from the peak fields. The radial fields indicate that the cavity will provide substantial focusing (7.5 MV/m) as the beam passes through the entrance nose cone with similar defocusing fields at the exit nose cone. The defocusing at the end of the cavity should be mitigated by the beam increasing its energy by about 2.61 MeV in passing through the booster.

To mitigate the field dip in the “reflected cavity,” the cavity was shortened by 5 cm. After shortening the cavity, the fundamental frequency decreased to 485.53 MHz. To tune the cavity back to the desired fundamental, the cavity radius was decreased by 6.9 mm from the starting 12 cm radius, resulting in a fundamental frequency of 500.14 MHz. Plotting the same on-axis and 1 cm off-axis electric fields, shown in Figure 85, we see success in decreasing the dip in E_z to only 75%

of the peak fields. Peak E_z has also increased to 22.2 MV/m in the on-axis case. For off-axis fields, peak values increased for E_z and E_r as well, to 24.0 MV/m and 7.9 MV/m respectively. Unfortunately, shortening the cavity results in a decreased integrated gap voltage of about 2.27 MV (a decrease of 13% over the longer cavity) and increased increased radial electric fields. It is unclear at this point whether it is worth mitigating the dip at the expense of beam energy.

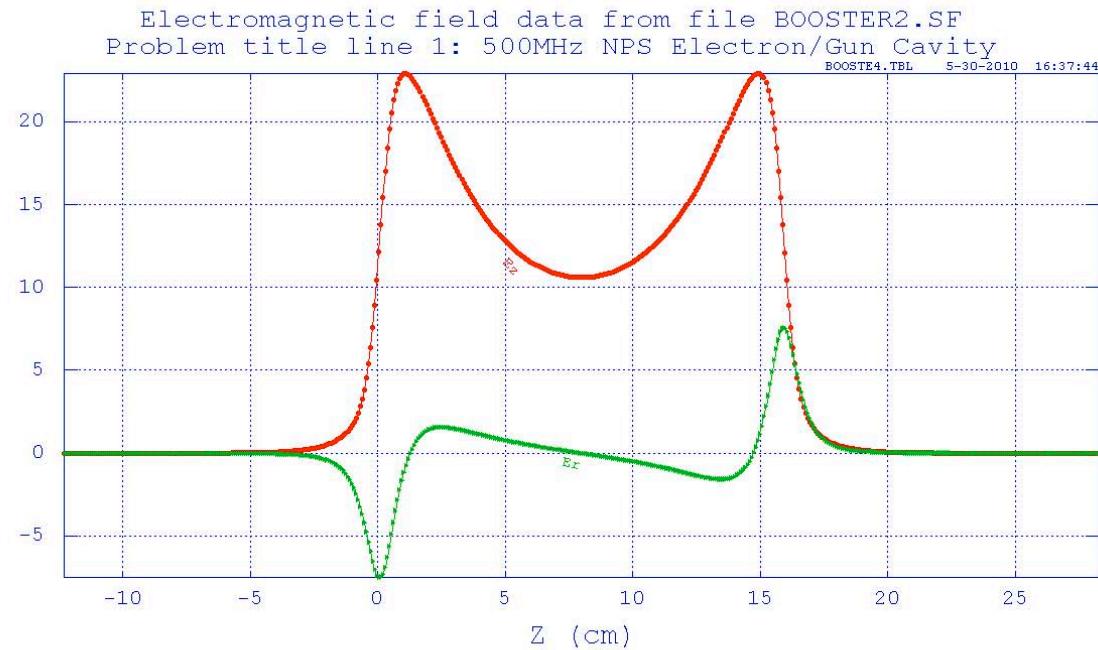
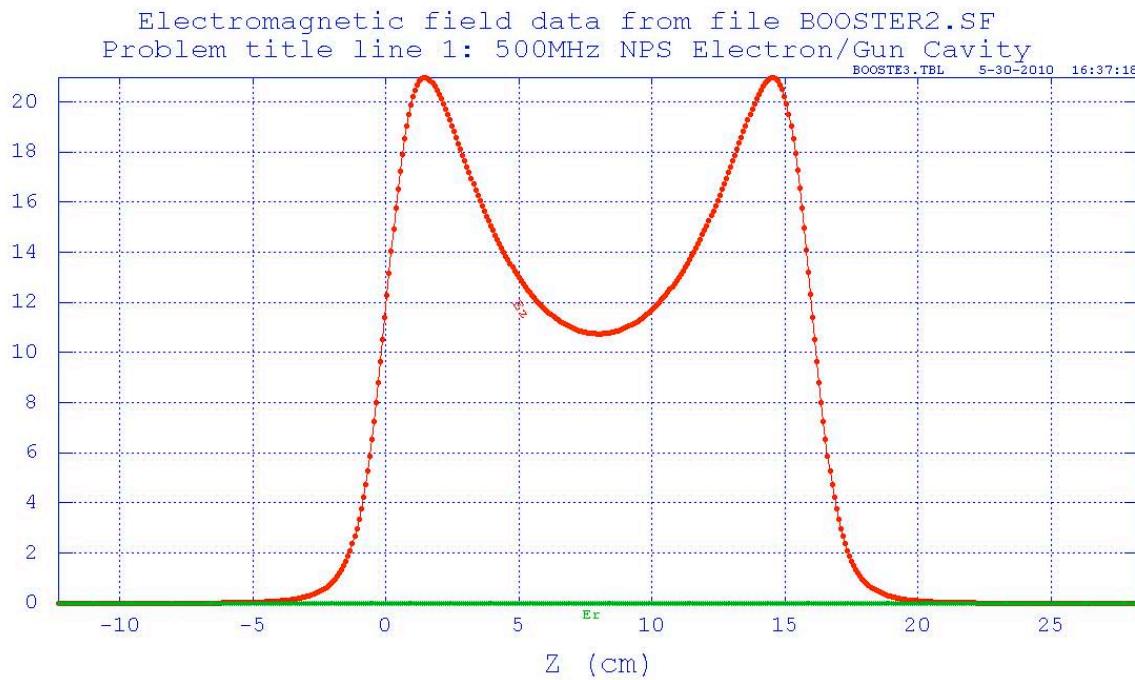


Figure 84. Electric fields along the electric axis of the long, double QW booster cavity (top) and a line displaced 1 cm radially (bottom). The red line is E_z and the green line is E_r .

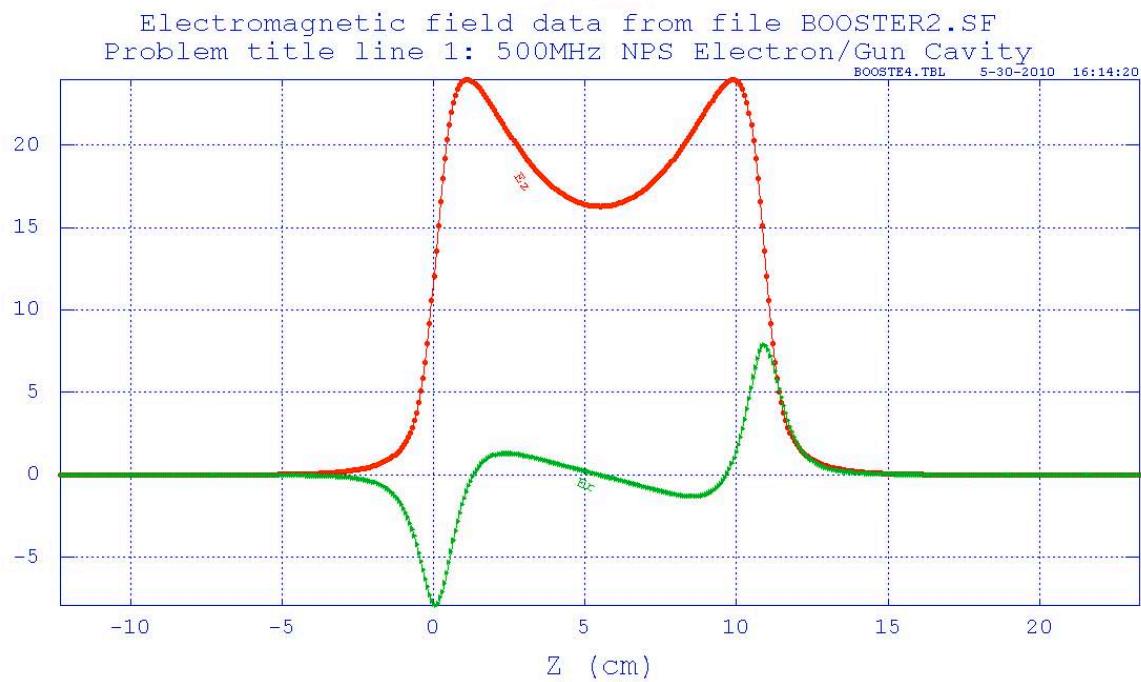
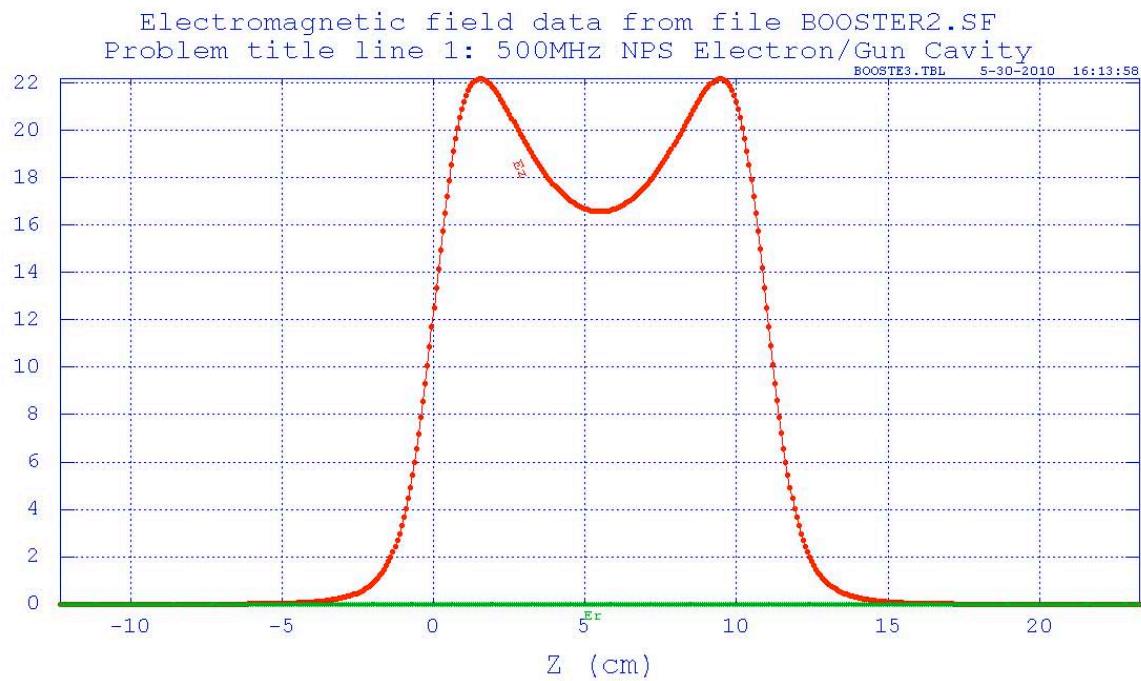


Figure 85. Electric fields along the electric axis of the short, double QW booster cavity (top) and a line displaced 1 cm radially (bottom). The red line is E_z and the green line is E_r .

b. Beam Modeling

With satisfactory cavity field maps, both versions were imported into GPT to determine their performance on a beam generated by the NPS gun. Both systems were started with a uniform 1 nC electron bunch 40 ps in length and 5 mm in radius. The NPS gun cathode is placed at a 6.5 mm retracted position and the beam launched at the optimal phase for spot size or emittance depending on the booster optimization desired. The gun cavity and booster cavities are kept at maximum field level. The two booster cavities will be compared based on phase acceptance, final beam kinetic energy, bunch charge preservation, spot size at the notional merger entrance, and emittance preservation.

We begin with the acceptance plots of the two booster cavities. After optimizing the gun for transport for spot size, the superconducting solenoid is optimized to force the beam to a waist at the exit plane of the booster cavity. Figures 86-89 show the example acceptance surfaces of the two cavities. In this case, 2D parameter surfaces are required for visualization as the phase delay and focus location from the gun's superconducting solenoid are tightly coupled.

For the long booster cavity, the emittance optimized acceptance plots are shown in Figures 86 and 87, where we have truncated the search volume to a small launch phase and gun solenoid strength (parameterized from 0 to 1) space where transport is expected. From the energy plot, we see that for optimal energy deposition into the beam, we should operate the booster at a phase delay of about 105° . The range of phase delays chosen ($50^\circ \leq \phi \leq 150^\circ$) all show excellent transmission as seen in the right hand plots. For the parameter space plotted, the gun solenoid has the most significant effect, with transmission decreasing rapidly above 0.38 (180 mT) and below 0.18 (86 mT). The lower left emittance surface seems to indicate that the emittance improves as the phase delay increases and the solenoid value decreases relative to the range considered.

When we look at the acceptance surfaces (spot size optimization) for the short booster cavity in Figures 88 and 89, we see that the effects of booster delay phase and solenoid focusing from the gun solenoid are fairly decoupled. Note that this plot is also only over a region of nearly full charge transport for purposes of optimization. If one considers the full range of phase delays, the surface becomes much more complicated, but as we are interested in transporting our beam at the greatest possible energy, we ignore these other areas. From these plots, we see that the booster should operate with a delay of about 100° from the cavity launch phase and the gun solenoid should have a peak field of about 150 mT. In general, both booster designs have equivalent phase delay acceptance ranges of $50^\circ \leq \phi \leq 150^\circ$ with respect to the gun launch phase and gun solenoid peak fields of between 100-200 mT being satisfactory for either optimization scheme.

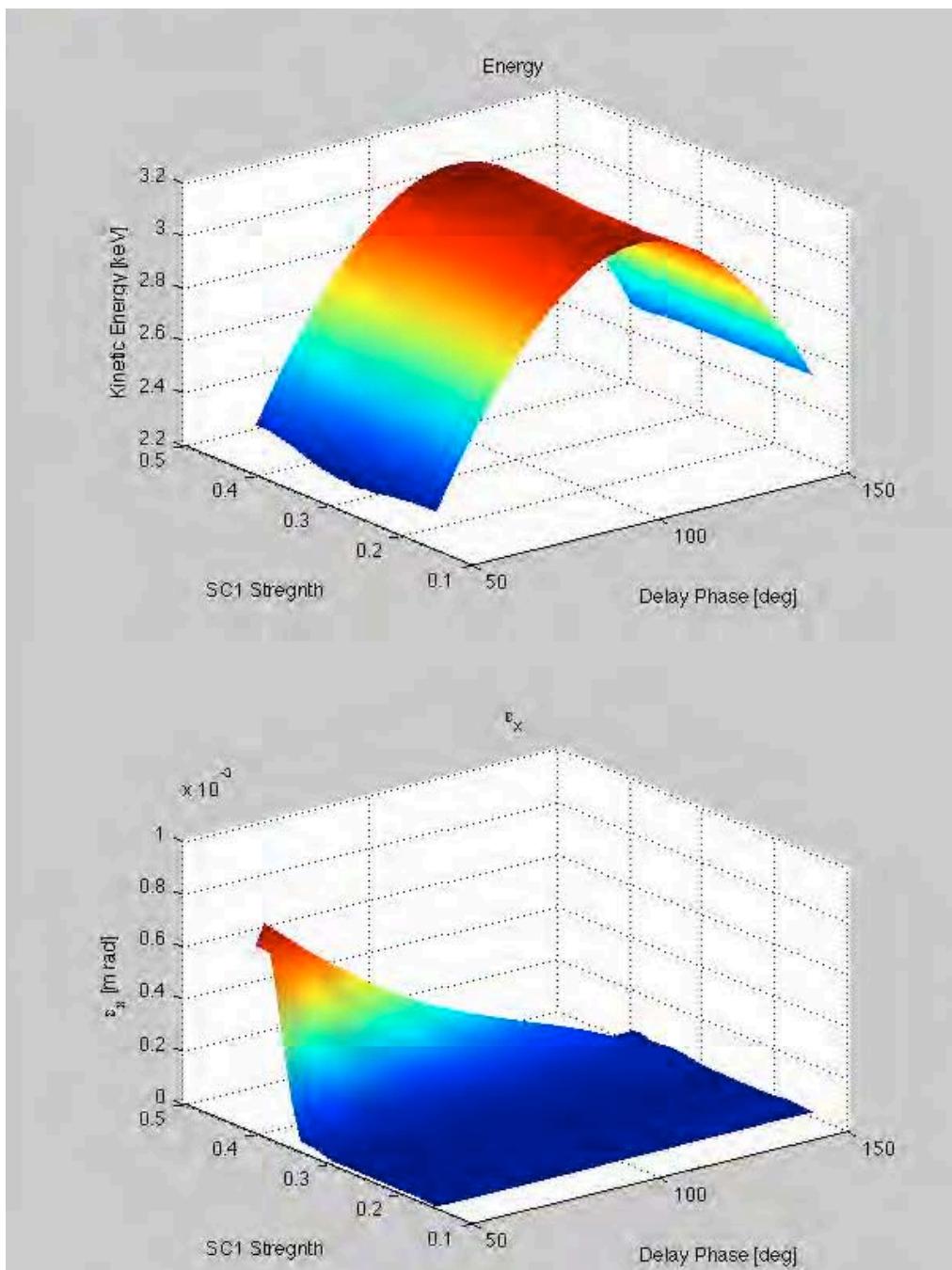


Figure 86. Acceptance plots for the long, double QW booster cavity, optimized for smallest transverse emittance at the booster exit, are shown. The upper plot shows average kinetic energy of the beam as the cavity phase delay is changed. The bottom plot shows a strong emittance dependence on solenoid strength for higher solenoid strengths.

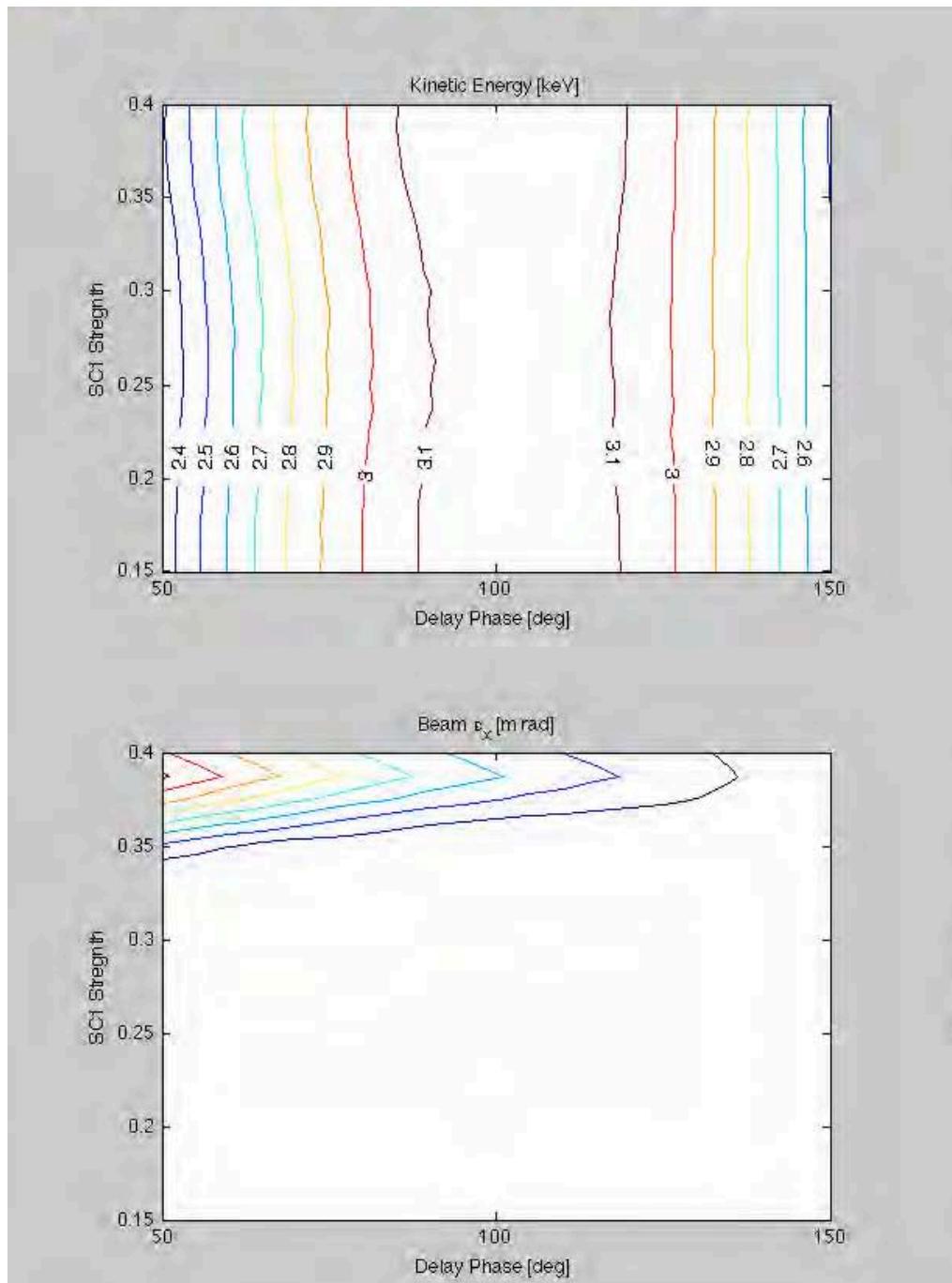


Figure 87. Projections of the plots shown in Figure 86.

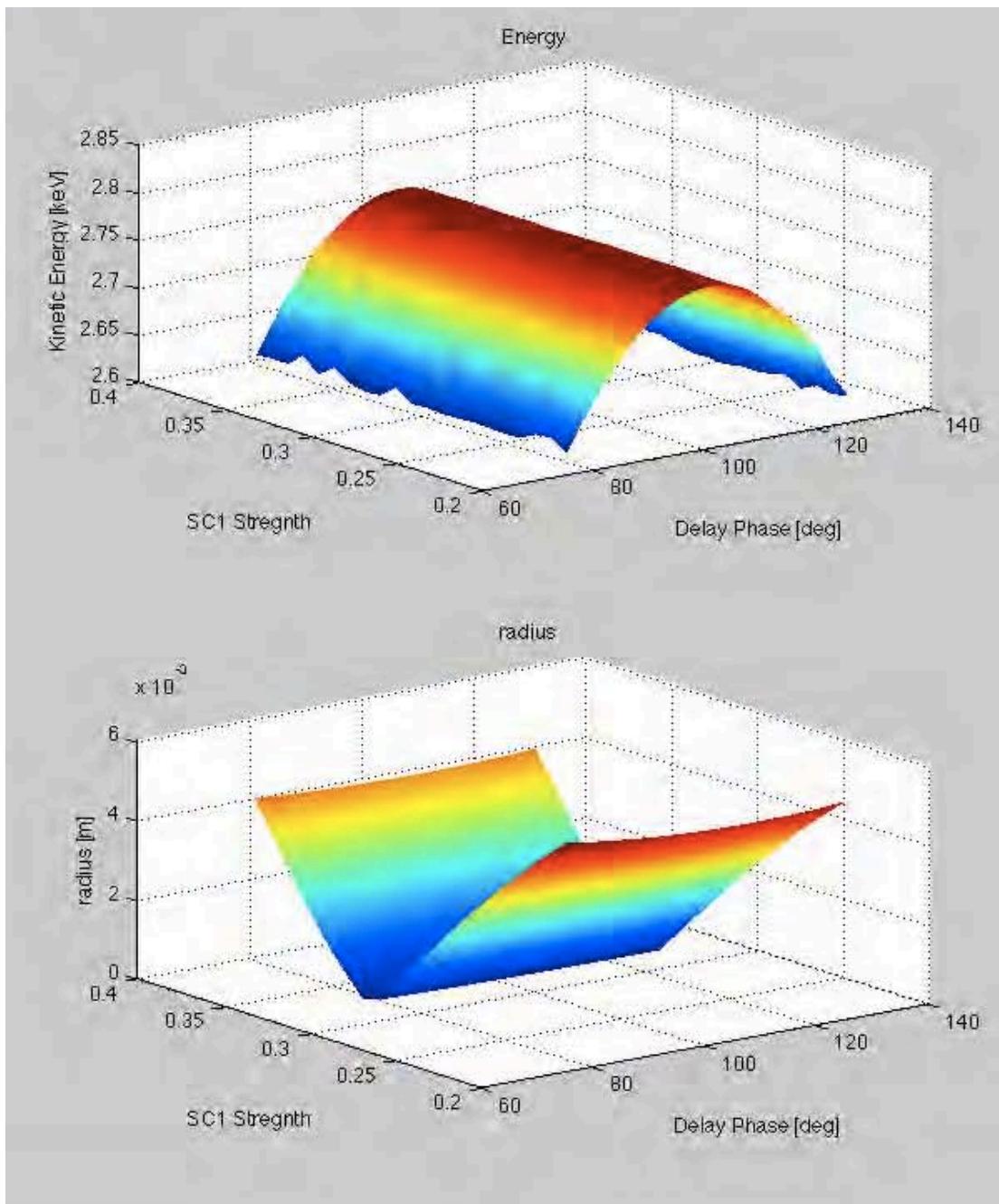


Figure 88. Acceptance plots for the short, double QW booster cavity, optimized for smallest spot size at the booster exit, are shown. The upper plot shows the average kinetic energy of the beam as the cavity phase delay is changed. The bottom plot shows a strong spot size dependence on solenoid strength.

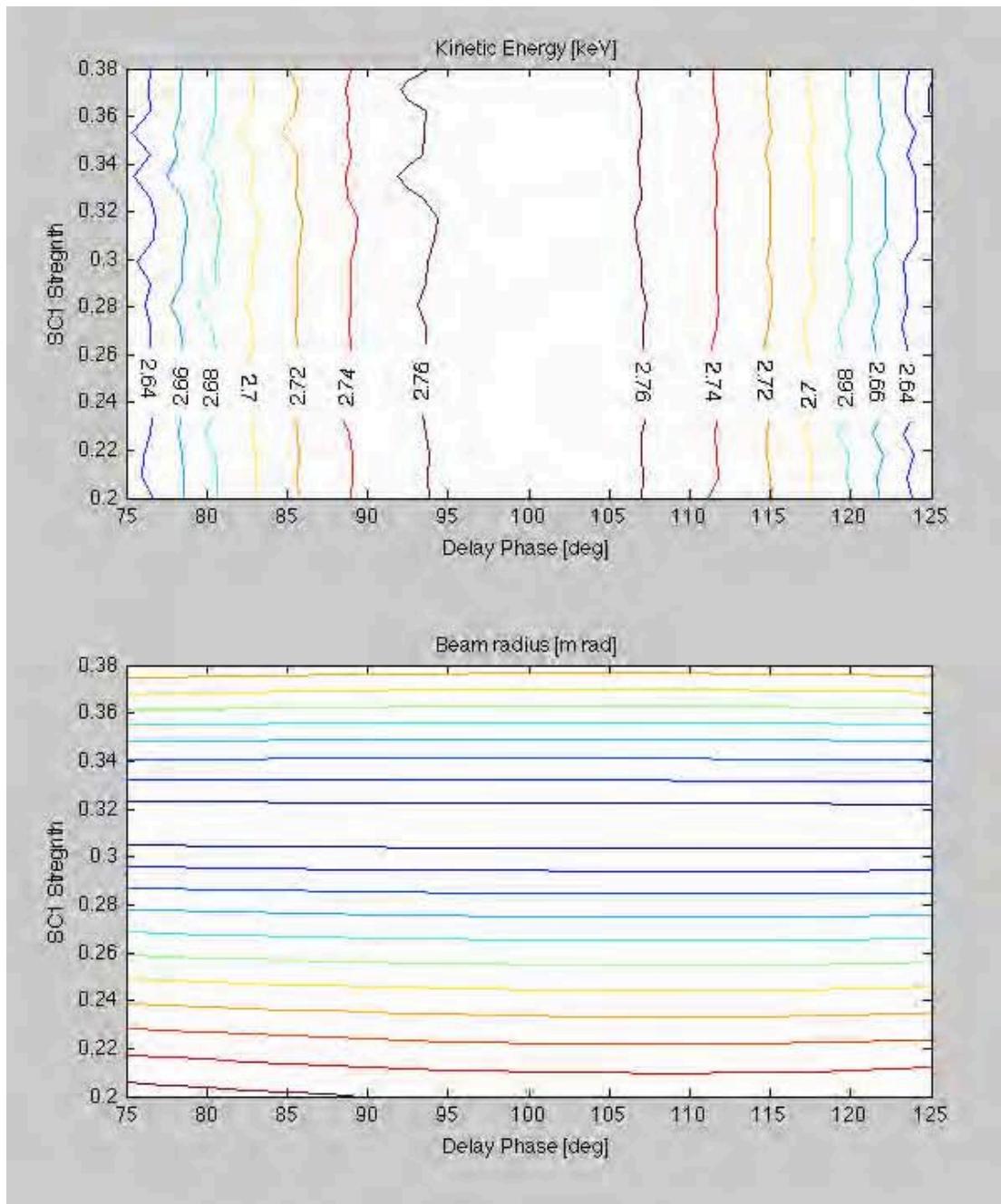


Figure 89. Projections of the plots shown in Figure 88.

If we then continue the transport optimizations, we can obtain a final system for comparison to the pure gun and transport cases considered earlier. Figure 90 shows the beam envelopes for the spot size optimized booster designs. With the compact nature of this booster configuration and assuming an additional 0.5 m from the booster solenoid for the cryomodule containment, the injector cryomodule would terminate at about the $z = 1.3$ m. We allow for an additional 1.7 m propagation distance similar to the propagation distance in the gun design to the normal conducting solenoid. We see that both the long and short booster cavities have an average radius of about 4 mm at the termination point, although both suffer from halo generation due to lower energy particles ejected at the beam focus. However, due to the disparity between the average and maximum radius, the halo charge makes up little of the total bunch charge for both designs. Both designs achieved full transport of 1 nC for all optimizations. We also see in Figure 90 that the long booster results in a higher beam energy, by about 210 keV, which is about half of what was expected from the integral fields. Both cavities, when optimized, apparently run slight off crest. Both the long and short booster designs fall about 10% short of their maximum energies, but still above the desired design goal of a 3 MeV beam at the exit, when optimized for the smallest spot size.

Figure 91 shows the transverse emittance evolution for the injector designs. For both cases, the gun and booster cavity parameters were optimized to obtain the smallest transverse emittance at the termination point. Neither cavity significantly improves over the emittance values found from just the gun with both designs having final emittance values of 20-30 μm . The short booster cavity does appear to have an advantage over the longer one as it does not add as much emittance in the process of acceleration. Of note is that neither booster greatly increased the transverse emittance, indicating they are still good candidates for boosters. Additionally, we see that optimizing the injector system for transverse emittance causes both booster cavities to run further off crest, resulting in a decrease in final beam

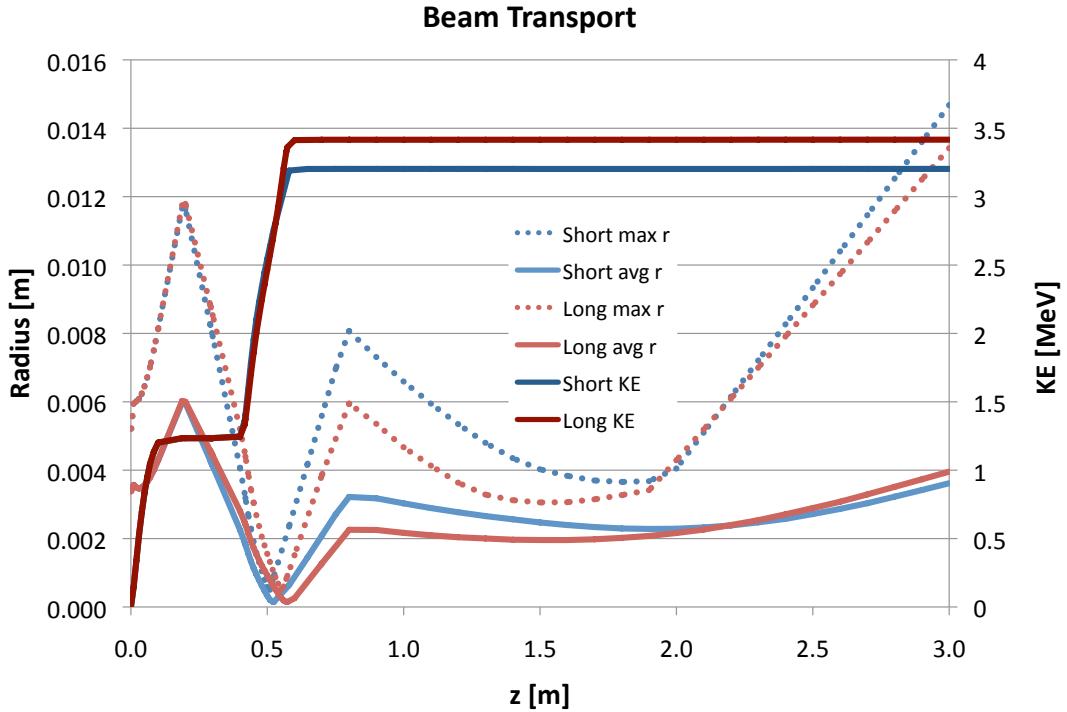


Figure 90. Beam transport envelopes for the long and short double QW booster cavity designs. The NPS gun is optimized for smallest spot size at the gun exit with the cathode 6.5 mm retracted for both designs.

energy (though still above the 3 MeV design goal) of about 100-200 keV with the longer booster cavity suffering more than the shorter cavity. If we compare these optimized emittance values with the transverse emittance values from the spot size optimizations ($\epsilon_{long} = 13.5 \mu\text{m}$, $\epsilon_{short} = 17.6 \mu\text{m}$) we see that there was no significant improvement over the spot size optimization. In fact, the long booster found a worse configuration for emittance.

The final comparison points between the original gun design and the proposed injector are in longitudinal emittance and energy spread. Figure 92 shows the output transverse emittance trace space for the injector system. From these plots, we can observe that there is significant energy spread and bunch length change. The spot size optimized beams show 20%-30% increase in bunch length, while the

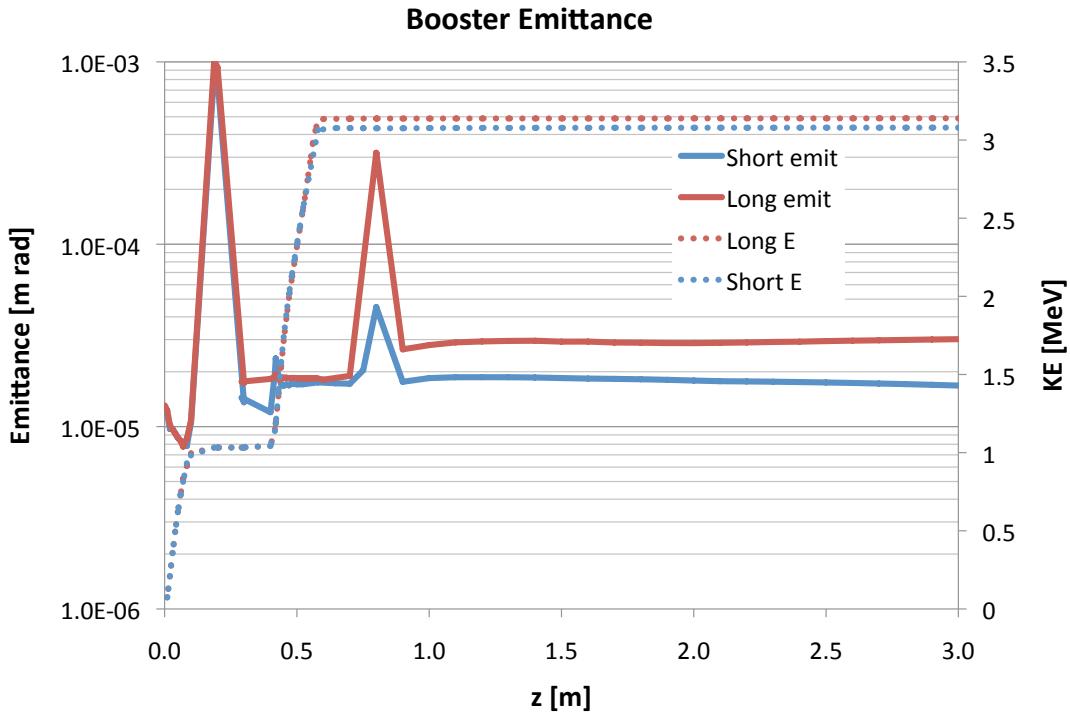


Figure 91. Beam emittance evolution for the long and short double, QW booster cavity designs. The NPS gun is optimized for best emittance at the gun exit with the cathode 6.5 mm retracted for both designs.

emittance optimized beams show a decrease in bunch length of about 10%. Table 5 shows the numerical values related to Figure 92 for the four injector and two gun optimization cases.

Table 5. Longitudinal emittances for the NPS gun and proposed injector designs.

Design	Optimized For	ϵ_z [eV s]	Energy Spread [%]
Gun	spot size	1.811×10^{-8}	1
Gun	emittance	1.054×10^{-8}	1
Long booster	spot size	1.764×10^{-7}	7
Long booster	emittance	7.672×10^{-8}	8
Short booster	spot size	1.688×10^{-7}	9
Short booster	emittance	7.053×10^{-8}	6

Recalling that the gun cavity parameters were optimized separately from the booster to isolate the booster contributions, it is definitely possible that a combined optimal configuration between the gun and booster could be found that would decrease the emittance values further. No clear “winner” was seen in this study, however with a higher output energy and comparable optimized spot size and emittances (transverse and longitudinal), the longer cavity may provide slightly greater benefit in an injector design.

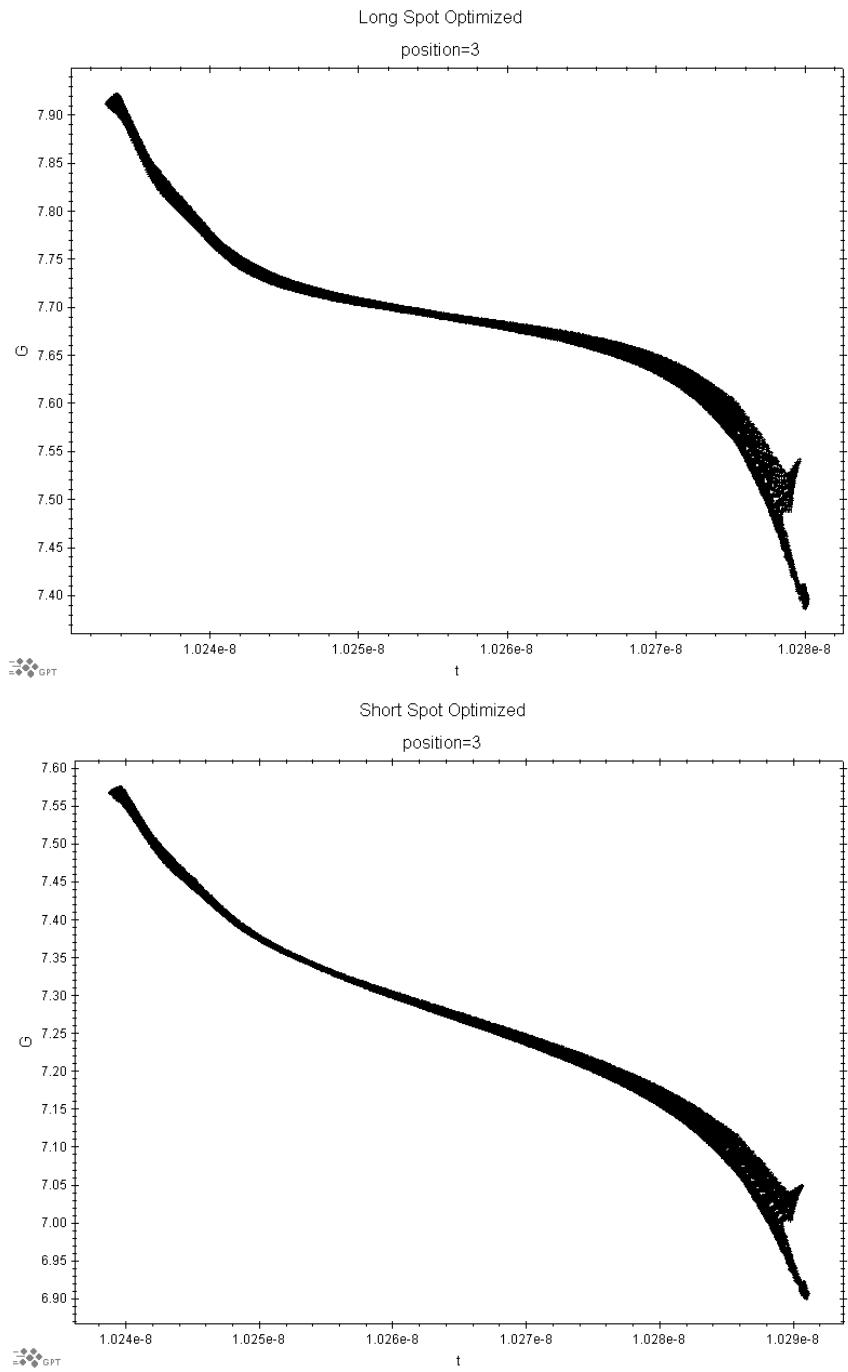


Figure 92. Longitudinal beam emittance plots for the long and short double QW booster cavity designs. The upper plots are for the long booster, and the lower plots are for the short booster.

B. SRF GUN/BOOSTER EXPERIMENTS

Simulation provides many insights into how systems work and can provide the experimentalist a platform upon which proposed experiments may be tried without fear of damaging expensive equipment. Simulations also provide the researcher the opportunity to explore larger data sets than possible with actual experiments. However, there is no substitute for developing hard data by experimentation with a physical system. If the experimental results compare favorably with simulation data, it provides confidence in both—that the simulation parameters are representative of the physical system and that the experiment is behaving as predicted by simulation.

As part of this philosophy, we have already presented simulation results for many possible experiments with the NPS gun/booster. In this portion of the chapter, we will discuss the actual experiments conducted during the development and construction of the gun/booster as well as the experimental results from the first beam testing of the first SRF gun in the United States [42].

1. Resonant Modes

The gun/booster system is fundamentally a resonant cavity for storing RF energy and coupling that energy into a beam of electrons. We previously discussed how the cavity shape was adjusted relative to the maximum fields in the cavity, but the situation is actually a bit more complicated. Every cavity will be able to support a sequence of discrete frequencies of RF, usually associated with higher-order modes. For a cavity intended to operate at the fundamental frequency, such as the NPS gun/booster, it is desirable to have any other higher-order modes be well separated from the fundamental frequency.

The resonant frequency is one of the first characteristics that can easily be tested, it was also one of the first to be simulated. Two separate software packages were used to simulate the cavity and determine the resonant modes: Superfish and COMSOL Multiphysics.

a. Superfish Model

To evaluate the cavity and the separation of resonant frequencies, we started with Superfish. Recall from the previous section that Superfish uses the meshed area to calculate a fictitious magnetic current density, \mathbf{K} , and solves for the frequencies at which the driving magnetic current required to maintain the electromagnetic fields goes to zero. This is captured in the equation

$$D(k^2) = kc \frac{\int \mathbf{H} \cdot \mathbf{K} \, dV}{\int \epsilon (\mathbf{H} \cdot \mathbf{H}) \, dV}$$

introduced previously. As the frequency, and thus the wavenumber k , is changed over the range provided, Superfish looks for locations of zero crossings of $D(k^2)$ and where the first derivative with respect to k^2 is -1. These locations correlate to cavity eigenfrequencies and the resonant frequencies are then provided to the user in an output file.

b. COMSOL Model

COMSOL Multiphysics is a finite-element code designed specifically to couple multiple physics areas together into single simulations. It consists of a core set of solvers and physics libraries with the ability to add additional libraries as the user requires. We used COMSOL primarily as a validation tool for the model as it has the ability to import CAD models directly. After obtaining the CAD model from Niowave, it was imported and established in the COMSOL environment. For the modal analysis and field solving performed, we used the electromagnetic waves package (called TM Wave) in the RF module and the dedicated eigenfrequency solver within that package to find the first 10 lowest eigenfrequencies. The COMSOL field map shown in Figure 93 for the fundamental mode, confirmed the locations of the peak electric and magnetic fields predicted by Superfish.

c. Comparison with Experiment

To measure the resonant modes of the cavity, the machined pieces of the cavity were assembled prior to welding. To ensure good conductive contact at

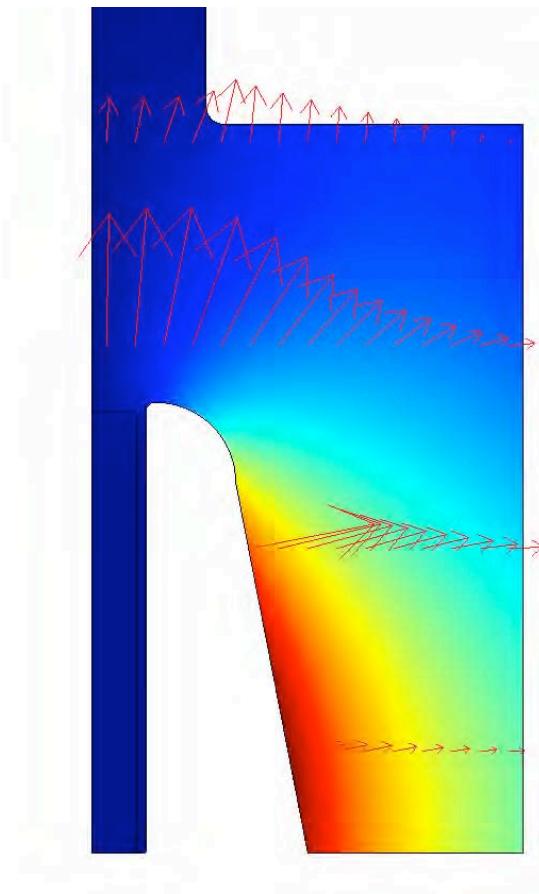


Figure 93. COMSOL field map of the simplified NPS SRF booster showing relative electric field strength and direction (arrows) and relative magnetic field strength (red, strongest → blue, weakest).

the joints between pieces, weights were placed on top of the cavity. Figure 94 shows the cavity as it was tested for resonant modes. An RF antenna was placed near the beam exit as the transmitting antenna and a receiver was located near the nose cone (this connection is not visible in the figure). A network analyzer was used to scan across frequencies looking for strong signals on the receiver lead.

The network analyzer scans across a user determined window of frequencies and measures the signal loss between the transmitted and received signal. For frequencies where the cavity is not resonant, losses are high. When a resonance is

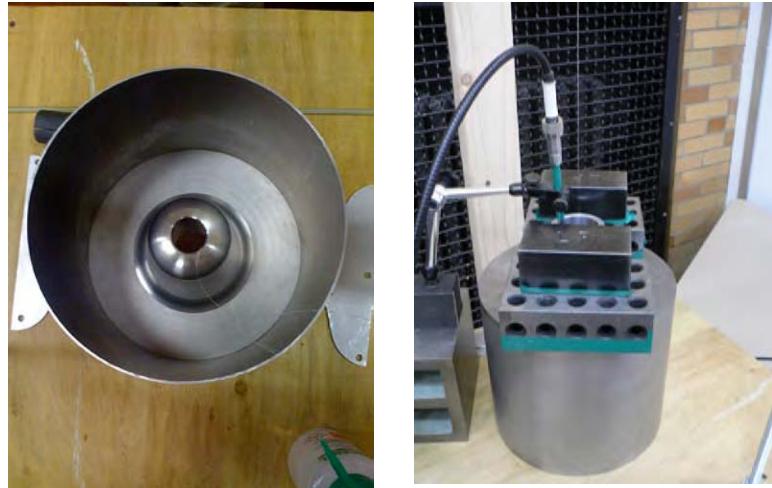


Figure 94. NPS gun/booster cavity configuration for eigenfrequency testing. On the left, is the cavity with the output end removed to see the nose cone. On the right, we see the cavity as tested. The blocks on top are there to provide weight to make good electrical contact with the sides of the cavity.

found, the output signal from the network analyzer spikes, as can be seen in Figure 95. If one zooms in on one of the spikes in the wideband sweep, the classic resonance plot appears. As we see in Figure 95, prior to welding, the cavity's fundamental frequency is 499.55 MHz. This is about 450 kHz lower than design, but the cavity will shrink slightly in the welding process and will contract when taken to cryogenic temperatures. Both processes will increase the resonant frequency. The data collected from this experiment and comparisons to predicted resonant modes is reported in Table 6.

Reviewing the results compiled in Table 6, we see that there is excellent agreement between the two simulation codes and the actual measured cavity frequencies. The overestimation seen in the fundamental frequency can be traced to the fact that the simulations use the ideal dimensions provided to the manufacturers. The measurements are taken on an unwelded, warm cavity, and the fundamental frequency can be expected to increase after welding and when the cavity is cooled. Through the 7th eigenfrequency, the agreement between cavity and simulations continued to be



Figure 95. Screen captures of the network analyzer data from the NPS cavity in the booster configuration prior to welding. On the left is the fundamental resonance at 499.55 MHz. On the right is a wide band sweep looking for higher order modes.

exceptional. Frequencies above 3 GHz were unable to be measured with the frequency analyzer on hand at the time of the experiment.

Table 6. NPS gun/booster cavity eigenfrequencies and comparison with simulation

Eigenfrequency Number	COMSOL [MHz]	Superfish [MHz]	Measured [MHz]
1	501.18	500.85	499.55
2	1,091	1,091	1,089
3	1,597	1,597	1,596
4	2,107	2,110	2,103
5	2,271	2,270	2,272
6	2,405	2,402	2,402
7	2,629	2,628	2,623
8	3,034	2,988	
9	3,073		
10	3,501		

2. Bead Pull

A bead pull experiment is a method for mapping the electric fields of a cavity. The basic process is to pass a small conducting bead through the interior of the operating cavity. As the conductor passes through the fields of the cavity, it expels the fields, slightly changing the cavity phase and resonant frequency. From this phase shift, one can calculate the electric field as a function of position along a line through the cavity.

As we are most interested in the on-axis fields, a thin fishing line was fed through the center of the cavity using the setup shown in Figure 96. The line was connected to a small electric motor attached to a spool for adjusting the bead's position in the cavity. The bead is a small piece of stainless steel hypodermic needle threaded onto the fishing line and held in place with knots. A weight is attached to the free end of the line to maintain constant tension on the line via a pulley system that also ensures the line passes through the cavity as close as possible to the electric axis. Using a ruler, adjustments are made in the bead position within the cavity and the received RF amplitude and phase shift were measured.

Using the motor, the bead was pulled straight through the cavity from exit to nose cone and the phase shift recorded as a function of time. The start and end positions were noted over a total transit distance of 20 cm and recorded for later use. We determine the field profile from the frequency shift as a function of phase shift recorded at a specific bead position [43], [44],

$$\frac{\delta f}{f_0} = \frac{\delta\phi}{2Q_0}, \quad (\text{IV.19})$$

where δf is the frequency shift, f_0 is the fundamental frequency of the cavity, Q_0 is the cavity quality factor, and $\delta\phi$ is the phase shift observed. We can relate this to the fields at the bead position [44] through the equation

$$\frac{|E|^2}{U} = \frac{\delta f}{f_0} \frac{1}{\epsilon_0 \pi R^3}, \quad (\text{IV.20})$$



Figure 96. LCDR Sean Niles stands next to the measurement system for the cavity bead pull. A transmitting antenna is inserted from the top of the cavity and a receiver below. The pulleys are connected via thin fishing line and attached on one end to an electric motor and a weight on the other to keep the line taught and to provide a reference point for linear position measurements.

where $|E|$ is the magnitude of the electric field (only E_z on axis for this cavity and frequency), U is the stored energy in the cavity (assumed constant as the perturbation due to the bead is assumed small), and R is the radius of the bead. Combining equations (IV.19) and (IV.20), we find

$$\begin{aligned}
 \frac{|E|^2}{U} &= \frac{\delta\phi}{2Q} \frac{1}{\epsilon_0 \pi R^3} \\
 |E|^2 &= \frac{U}{2Q\pi\epsilon_0 R^3} \delta\phi \\
 |E| &\propto \sqrt{\delta\phi} .
 \end{aligned} \tag{IV.21}$$

Taking the data displayed in Figure 97, the phase offset is subtracted to obtain a true zero. We also use the magnitude of the phase shift to avoid imaginary values when applying equation (IV.21).

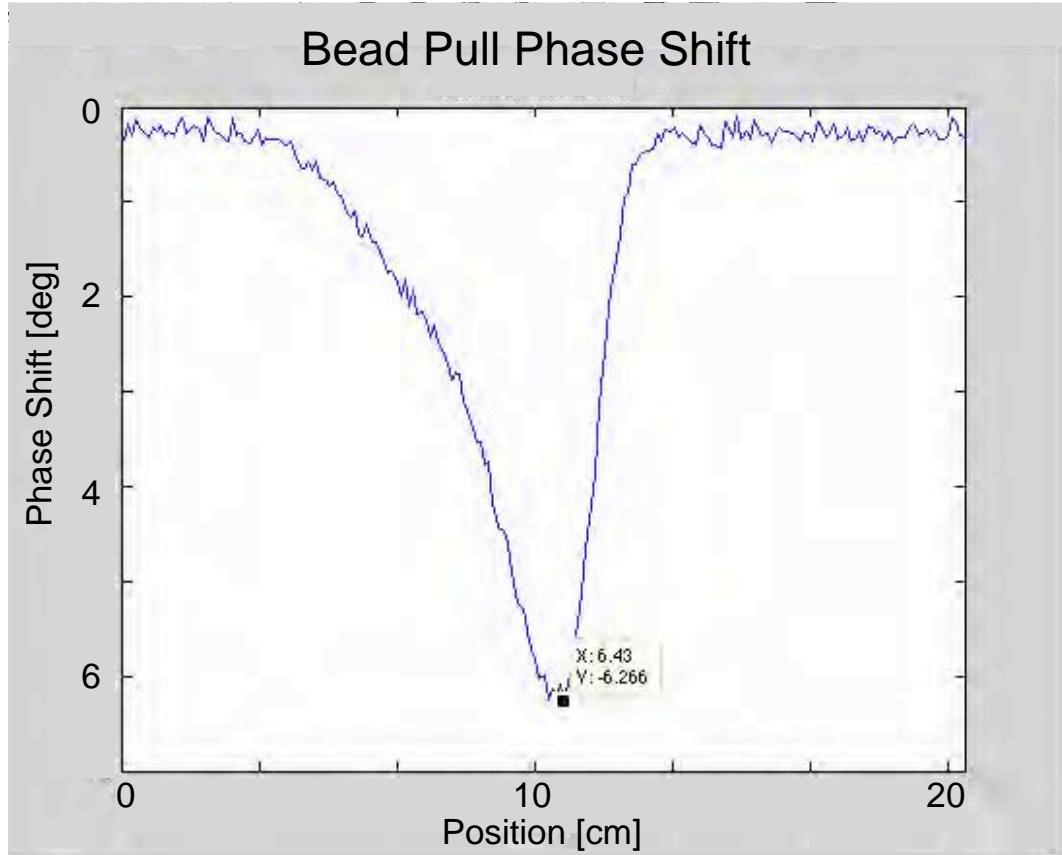


Figure 97. Phase shift as a function of position of the bead within the NPS booster cavity. The pull required 12.248 s to traverse 20 cm of the cavity and beam tubes.

To make a meaningful comparison with the field profile from Superfish, a plot of E_z is obtained along the cavity axis (in the booster configuration) and normalized to the peak field value. The peaks of the Superfish plot and the normalize peak of the phase data plot are aligned and the relative positions of the phase plot adjusted (time reversed since we measured from exit to nose cone) in order to compare the two

data sets. The resulting plots are plotted in Figure 98. From this plot, the bead pull profile shows good agreement with the Superfish profile.

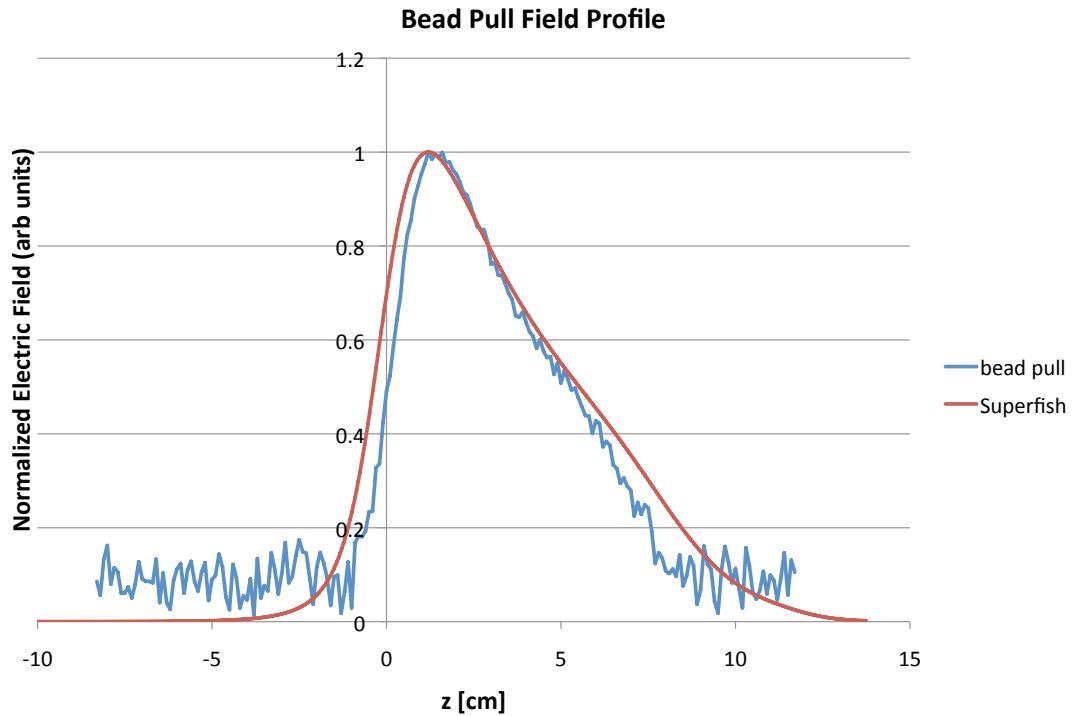


Figure 98. The electric field profile generated from a bead pull of the NPS booster cavity. The normalized bead pull plot is compared to a normalized plot generated in Superfish.

3. SC Solenoid

A superconducting solenoid is installed just after the cavity to provide beam focusing for electron beam transport and for emittance compensation. Due to the solenoid being located near the liquid helium dewar, it is chosen to be superconducting to allow for high amperage currents (and thus, high magnetic fields) without generating a measurable heat load to the cavity. The solenoid is an aluminum spool wound with niobium-titanium wires embedded in a copper matrix and backed with an iron backer to provide a return for the magnetic field line circuit. The solenoid has 2,236 turns and was designed for a maximum 10 A current resulting in a peak magnetic field of 476 mT (more than twice the anticipated necessary field for beam focusing).

a. *Superfish Model*

The first models were developed using Superfish to predict the field profile and maximum field strengths in the iron back and along the beam axis. The field maps shown in Figure 99 indicate that within the normal operating parameters of the solenoid ($I_{sol} \leq 5$ A), the iron backer should not saturate. Above about 5.5 A, saturation begins to be a minor concern. The field profile shape is as expected and desired. The field extent is sufficient to affect the SRF cavity when it transitions to superconducting, so we must remain vigilant to ensure no residual magnetic field remains on the solenoid to prevent cavity field degradation. Should the iron core become magnetized, this can be corrected by degaussing the solenoid prior to the next cavity transition. The flux trapping process will be discussed later in this section.

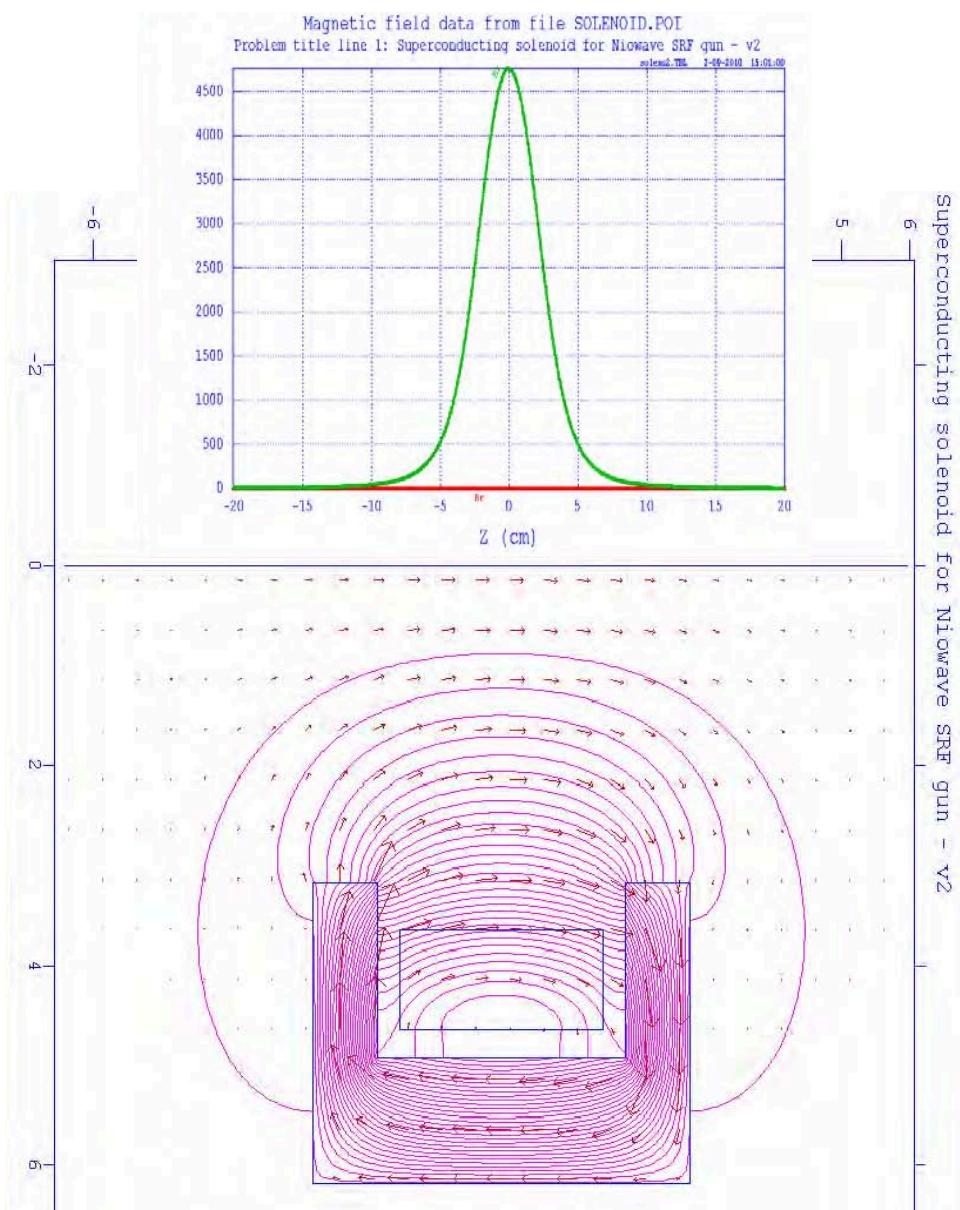


Figure 99. The superconducting solenoid field map generated by Superfish. The lower plot is a 2D map showing the magnetic field lines of the solenoid. The inset at the top is the magnetic field strength along a line through the magnetic axis of the solenoid. The peak magnetic field along this line occurs at the center of the solenoid and $B_{pk} = 4,760$ Gauss (476 mT).

b. COMSOL Model

Performing the same simulation in COMSOL provided us with confidence about the peak fields and saturation onset currents. Figure 100 shows the field map and on-axis magnetic field profile obtained from COMSOL. The results from COMSOL agree extremely well with the results obtained from Superfish.

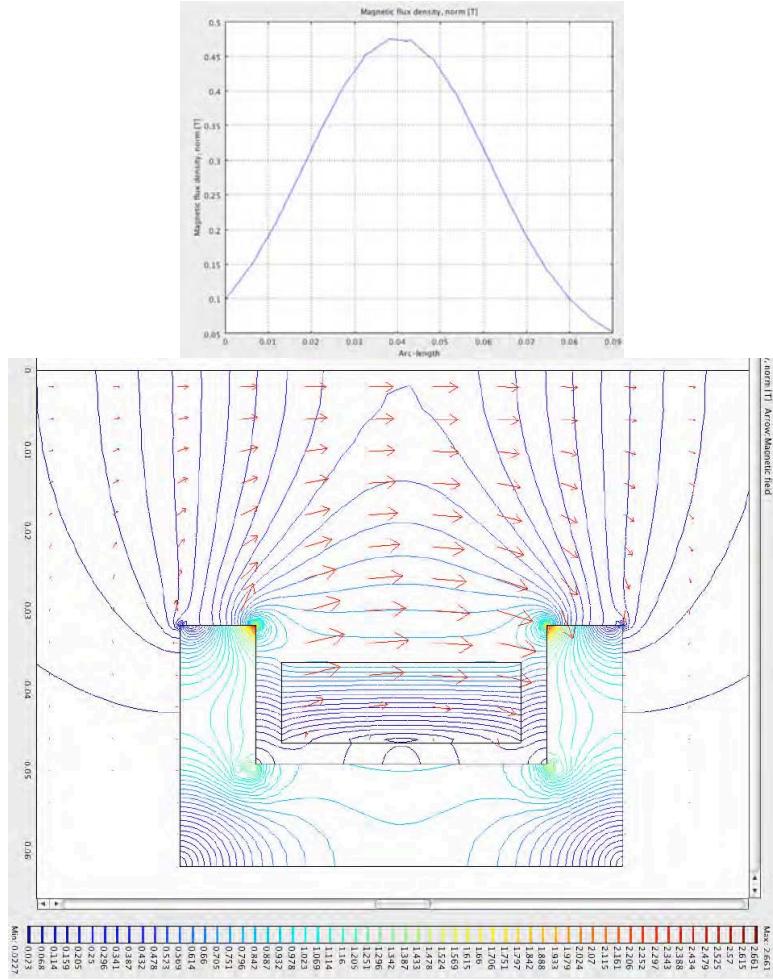


Figure 100. The superconducting solenoid field map generated by COMSOL. The lower plot is a 2D map showing that the maximum magnetic flux density (contours) is 2.66 T, occurring at the inner corner of the iron backer and the relative strength of the magnetic field (arrows). The inset at the top is the magnetic field strength along a line through the magnetic axis of the solenoid. The peak magnetic field along this line occurs at the center of the solenoid and $B_{pk} = 476$ mT.

c. Comparison with Experiment

To verify the design, the solenoid field was measured using a Hall probe in a tabletop test. This type probe has a Hall effect sensor at the tip of a long wand that measures magnetic field strength. The experimental measurement set up is shown in Figure 101. The hall probe is mounted in a nonmagnetic plastic block that is machined to fit into a U-channel aluminum bar. The U-channel and meter stick are passed through the solenoid and the block adjusted such that the probe will pass as near to the solenoid axis as possible. As the probe is slowly moved through the solenoid, magnetometer readings are taken. Measurements were made entering and exiting the solenoid from both sides to ensure any bias was averaged out.

Since the solenoid measurements were made at room temperature, we necessarily performed them at low current. There is no reason to believe that the field profile will change when the temperature is decreased and the current is increased where the solenoid is at superconducting temperatures. Figure 102 shows the experimental data compared with the appropriate plot generated from Superfish for three different applied currents (20 mA, 40 mA, and 60 mA). All plots show excellent agreement between simulation and experiment.

The solenoid was then placed in a cryogen bath to test its superconducting properties. As the solenoid was cooled, its internal resistance was measured using a standard 4-wire measurement. At room temperature, the solenoid had a resistance of $165\ \Omega$. Niobium-titanium has a transition temperature of about 15 K and we expect the resistance just prior to transition to be about $1.6\ \Omega$. As expected, the resistance dropped to $0.8\ \Omega$ as the sensing wire resistance steadily increased to $1\ k\Omega$ indicating the solenoid was superconducting. The solenoid was then slowly ramped to its design current of 10 A and operated there for at least a minute without quench. During the ramp up of current, the power supply voltage remained linear with the current increase indicating no significant resistive heating in the power leads. To test the solenoid at quench, we applied a 10 A current to the solenoid and then lifted it

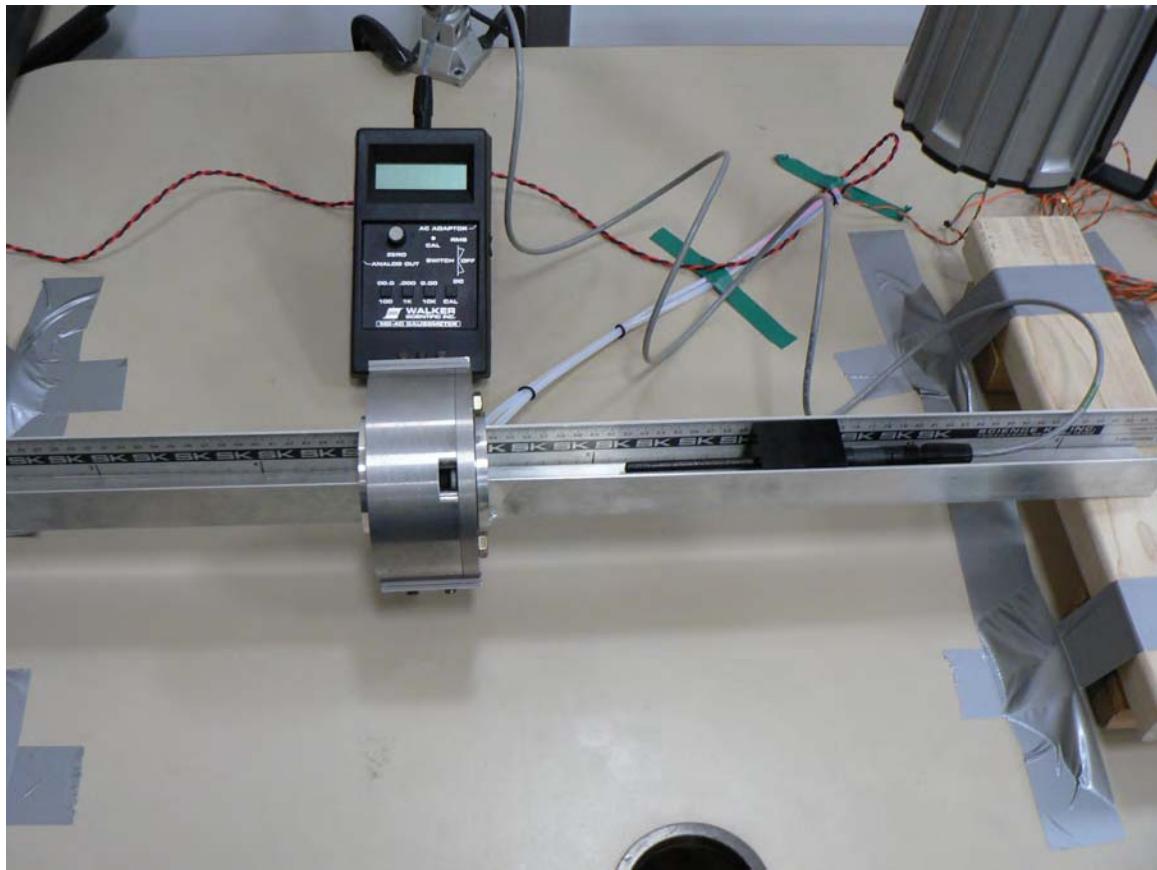


Figure 101. The experiment setup for measuring the magnetic field strength through the center of the superconducting solenoid.

from the liquid helium bath. As the temperature increased to transition, we saw no increase in the sensing line voltage until a large spike was encountered, indicating the solenoid was no longer superconducting. The solenoid was immersed back into the liquid helium bath, but due to resistive heating did not return to the superconducting state until current was removed to allow for the solenoid to cool. The solenoid was removed briefly, Figure 103 for inspection before returning to the bath to measure its inductance.

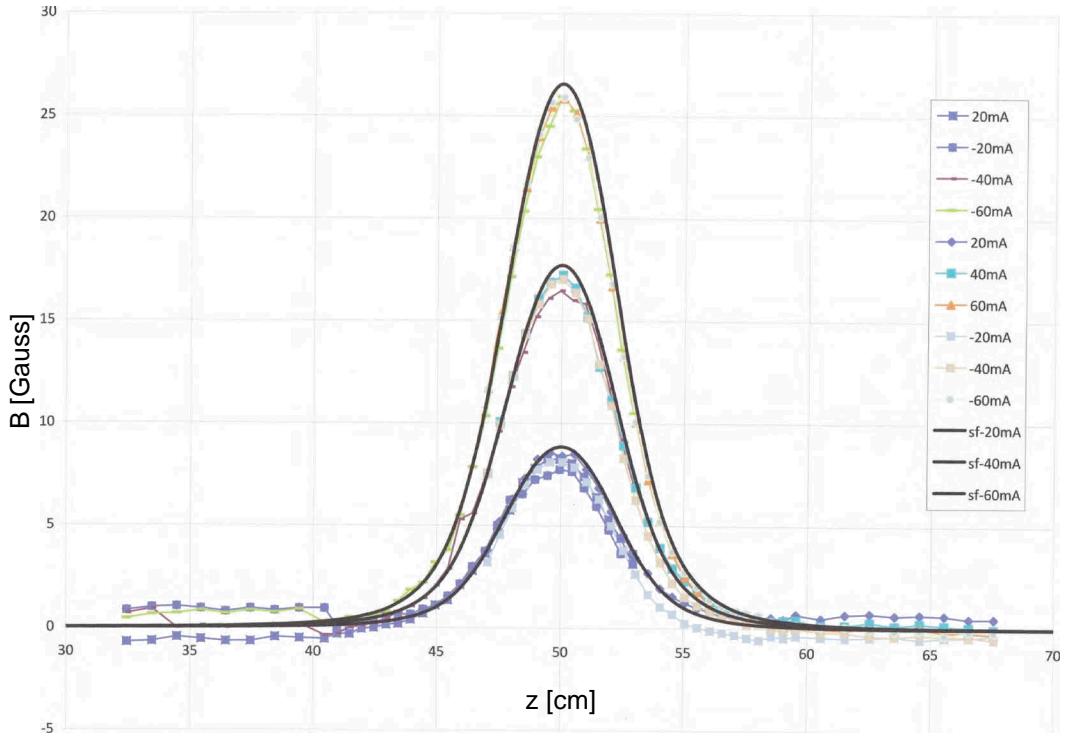


Figure 102. Experimentally determined axial magnetic field maps of the superconducting solenoid compared to Superfish models for various applied currents. Magnetic fields in this plot are measured in gauss.

To measure the inductance of the superconducting solenoid, we recall

$$V_s = \mathcal{L} \left(\frac{dI}{dt} \right) , \quad (\text{IV.22})$$

where V_s is the voltage across the solenoid and \mathcal{L} is the solenoid inductance. To obtain the current as a function of time, the power supply voltage is ramped in order to maintain a constant V_s . Differentiating and rearranging Ohm's law provides us with a method to calculate the time rate of change of the current as the applied voltage changes,

$$\frac{dI}{dt} = \frac{1}{\mathcal{R}} \frac{dV}{dt} . \quad (\text{IV.23})$$

Using the power supply voltage as a function of time recorded in Figure 104, we can calculate the solenoid inductance. The line resistance is $\mathcal{R} = 100 \text{ m}\Omega$ and $\Delta V / \Delta t =$



Figure 103. Inspecting the superconducting solenoid after removal from the liquid helium bath.

600 mV / 6.6 s, giving a rate of change in the current of 0.91 A/s. Putting this value into our inductance equation, we find the solenoid has an inductance of about 0.55 H.

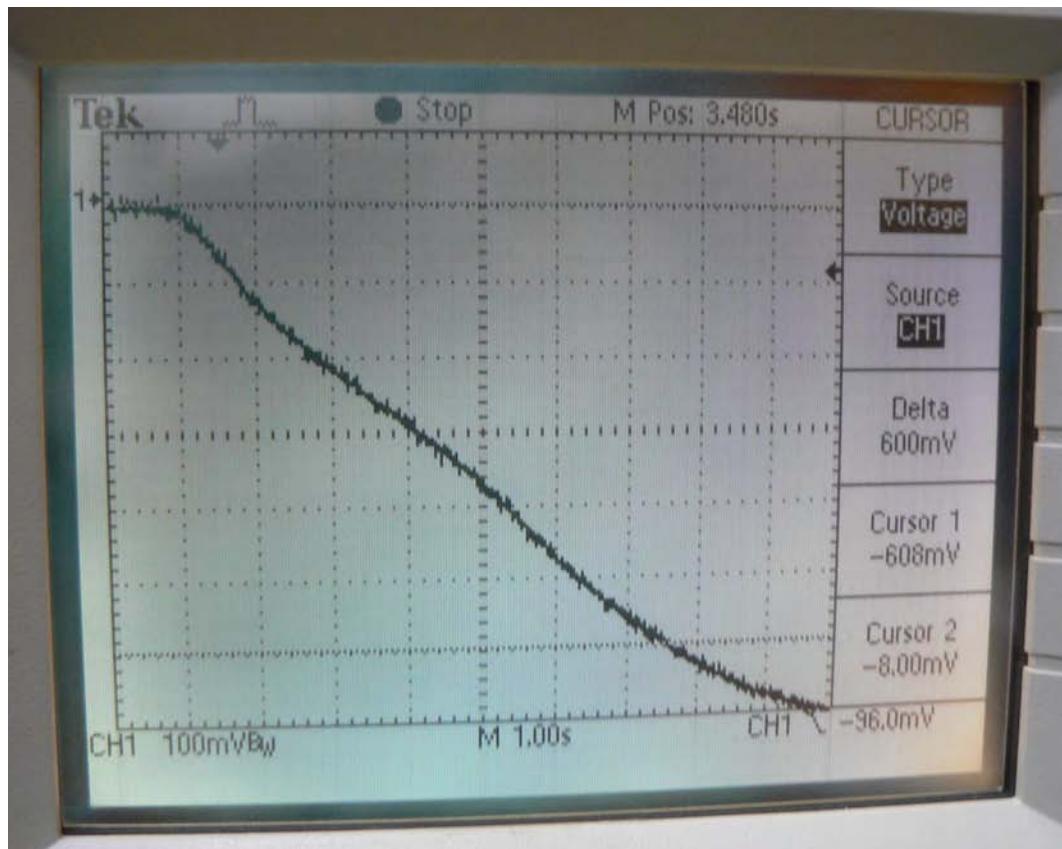


Figure 104. Power supply voltage as a function of time used to calculate the inductance of the superconducting solenoid.

d. Comparison with Simulation

For a solenoid to focus an electron beam, it must also cause the beam to rotate. This rotation is generated in the fringe fields of the beam, causing the beam to rotate as it enters the fields and then stop as it exits the fringe fields. The imparted transverse motion causes the beam to focus at some distant point before expanding again. Using GPT, we can model this motion and compare with an analytical calculation as a benchmark of both the simulations and our experimental measurements. To visualize the rotation of the beam, a uniform beam is broken up into four beamlets, each of which makes up 1/4 of the beam as seen in Figure 105. As the beam propagates, we observe that the fringe fields begin to rotate the beam (top right) and in the final picture, we see the output beam. Particle mixing has occurred during propagation due to internal beam Coulomb forces, but the lines between beamlets are still easily identifiable. From the simulation, we can estimate that the 1.2 MeV beam has rotated approximately 45° as it passed through the solenoid with $B_{pk} = 117$ mT.

We can perform a hand calculation to compare against what we observe in the simulation. The rotation angle experienced by a beam as it passes through a solenoid is related to the magnetic field strength, field extent, and beam rigidity through the relation [45]

$$2\theta = \frac{BL}{(B\rho)} , \quad (\text{IV.24})$$

where θ is the rotation angle, B is the magnetic field strength, L is the magnet's effective length, and $(B\rho)$ is the beam rigidity. The rigidity of an electron beam can be found by applying [46]

$$B\rho = 3.3356p \text{ [T m]} , \quad (\text{IV.25})$$

where p is the momentum and is given in GeV/c . To obtain the momentum, we must

first find γ from the total energy equation,

$$\begin{aligned}\gamma m_0 c^2 &= KE + m_0 c^2 \\ \gamma &= 1 + \frac{KE}{m_0 c^2},\end{aligned}\quad (\text{IV.26})$$

where m_0 is the rest mass of the electron. We know the kinetic energy of the beam is 1.2 MeV and can obtain the integrated field from the Superfish field maps we have generated (0.469 T m) at the stated peak magnetic field value. From these values, we calculate $\gamma = 3.35$ and can calculate $p = \gamma \beta m c = 1.63 \text{ MeV}/c$. Using these values and combining equations (IV.26) and (IV.25) and substituting into equation (IV.24), we have an expected rotation angle through this solenoid of

$$\theta = \frac{BL}{2(3.3356 p)} = 43^\circ, \quad (\text{IV.27})$$

which is in excellent agreement with the simulation.

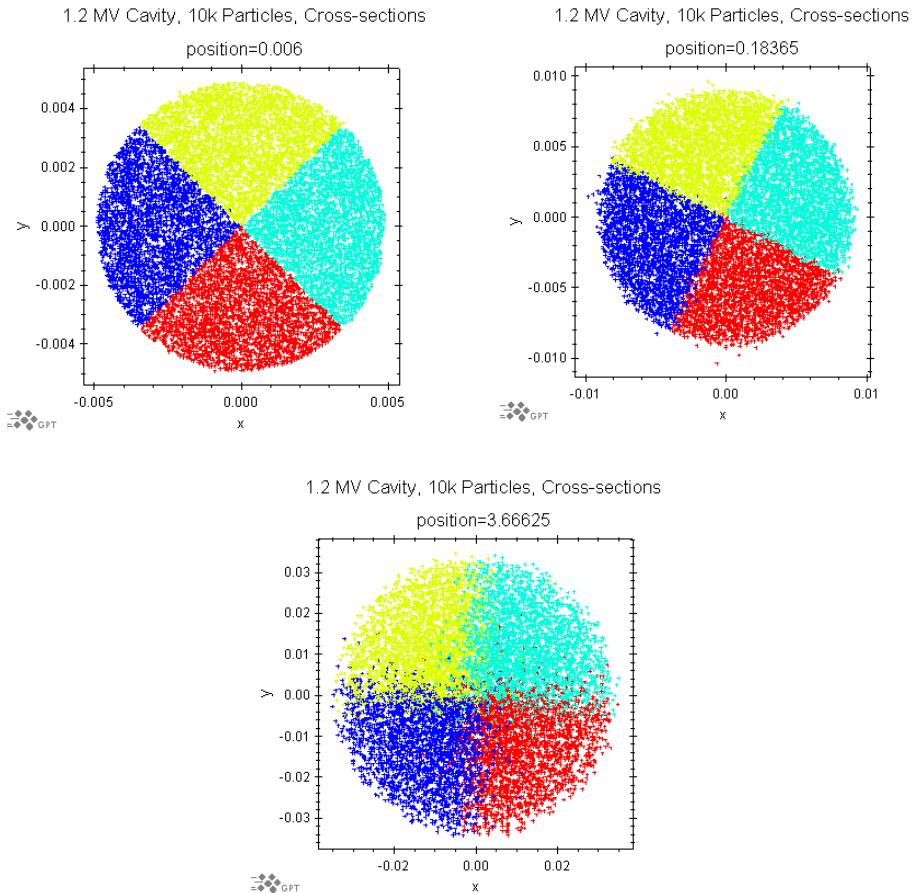


Figure 105. Cross sections of a 1.2 MeV beam from the NPS gun passing through the superconducting solenoid. Top left is the beam just after launch from the cathode. Top right is the beam at the center of the superconducting solenoid. Bottom is the beam after propagating to an output screen approximately 3.5 m from the solenoid.

e. *Thermal Transport*

Recall from the previous chapter that the superconducting solenoid is located outside the liquid helium dewar and thermally connected to the dewar via two solid, copper bus bars. It was hoped that this thermal connection would be sufficient to allow the solenoid to cool below T_c . However, during initial cold tests of the cavity, it was discovered that this cooling connection was insufficient to cool the solenoid fully. In normal operating mode, with no RF applied to the cavity, and no current applied to the solenoid, the coolest temperature, as measured on the outside of the solenoid case, was 11.3 K.

Because the solenoid is not superconducting, any current introduced into the solenoid will travel through the copper portion of the wire and cause resistive heating. This generated heat is tied to the liquid helium dewar and will cause an additional heat load to the helium bath, decreasing the operating time available on a single fill. Without opening the cryomodule, there is no way to increase the thermal connection of the solenoid to the helium dewar.

It is possible that there are additional external heat loads that are preventing the solenoid from cooling sufficiently and becoming superconducting. Our initial thought was that the coupler, which passes through the solenoid inside the beam tube, could be causing radiative heating since it is not cooled in any way. To test this hypothesis, we wrapped the vacuum flange where the coupler short connects to the beam tube to the off-gassing line of the cryomodule liquid nitrogen shield and insulated it. After an hour of applied cooling, no significant change in the solenoid temperature was seen. It is possible that the conductive heating from the beam tube could be the additional heat source we seek to mitigate, but this has not been tested to date.

To avoid introducing too much heat into the cryomodule with the solenoid, the current operating procedure is to run the solenoid in pulsed mode. The present gun status is such that the beam is of low enough energy that the fo-

ocusing fields generated with the reduced solenoid current are sufficient for present experiments. Solving this heat transfer problem is essential in the next generation of the NPS SRF gun.

f. Flux Trapping

Most physics students have at least a passing familiarity with the Meissner effect in passing—usually in the context of the simple demonstration of a magnet placed on a superconducting material that is taken through the transition temperature. As the superconductor’s temperature drops below T_c , the magnetic field from the magnet is expelled and the magnet levitates above the superconductor surface. The Meissner effect in a superconducting cavity behaves in a similar way. With the simple physics demonstration, both the magnet and superconductor are usually small and the effects of trapped flux are typically not observed. In a superconducting RF cavity, the raw amount of superconducting material makes the likelihood of impurities being present in the bulk material almost definite. These impurities can cause significant issues if the superconductor transitions in the presence of a magnetic field. Flux trapping is a condition where impurities or lattice defects in the superconductor act as trapping sites, resulting in only partial flux expulsion [12]. The pinned magnetic flux lines result in increased residual resistance increasing the cavity losses [47]. These increased losses result in a decreased cavity Q_0 and associated accelerating fields.

Since the NPS cavity design has a solenoid capable of developing a nontrivial magnetic field near a superconducting cavity, it is essential that we understand how this affects performance. As the solenoid is used, particularly at higher field levels, it is possible to induce a residual magnetic field in the solenoid when the current is removed. These field lines can pass through the walls of the cavity during transition, and, if not fully expelled, result in degraded cavity performance. There is no issue if the cavity is maintained superconducting as the magnetic fields from the solenoid will not be able to penetrate the cavity surfaces.

To test this effect and observe the impact on this particular cavity, we purposely attempted to induce flux trapping during the first cold tests of the cavity. In the first step, we degaussed the solenoid prior to cooling the cavity below T_c to ensure no residual fields from previous experiments remained. We then took the cavity superconducting and measured a cavity Q_0 value as a baseline. The solenoid was operated briefly at moderate current ($I = 4.2$ A) to magnetize it. After the cavity quenched, we refilled the helium dewar and brought the cavity back to the same operating state to measure another Q_0 value. Once back to the previous operating temperature, the Q_0 measurement indicates a moderate decrease in cavity performance. We then allowed the cavity to quench again and degaussed the solenoid to remove any residual magnetic fields. After cooling the cavity again, we were successful in returning to the original Q_0 value, indicating that the cavity can be affected by residual solenoid fields, but that the effect can be corrected by degaussing the solenoid when the cavity is normal conducting.

4. Cold Cavity Processing

When an accelerating cavity is first brought into service, it must be conditioned prior to full use. Conditioning typically involves putting the cavity into the expected operating environment (under vacuum, cryogenic temperatures, etc.) and placing it under an RF load. As RF is put into the cavity, any field emission sites will begin to emit and hopefully condition themselves away. Multipacting barriers may also be encountered as RF power levels change. Typically, once a cavity has been conditioned to a certain RF power level, unless something in the system changes (such as venting the vacuum), the cavity can be operated safely and easily below the maximum conditioned level. During conditioning, both field emission and multipacting are generating accelerated electrons that run into the walls of the cavity and beam tube and generate X-rays as they are stopped by the metal. By monitoring radiation levels, we can estimate the level of conditioning that takes place.

The NPS gun/booster cavity is designed to be a batch fill cavity: the cavity will not, as many other SRF cavities are, be cooled once to liquid helium temperatures and maintained there for long periods of time. As the cavity temperature is taken from room temperature down to 4 K, it must transition through a temperature region where the possibility of “Q-disease” becomes increased. Q-disease is the formation of hydride precipitates on the surface of the cavity [12]. These hydrides have much higher electrical resistance and contribute to localized heating of the cavity, possibly creating localized regions where the temperature of the cavity exceeds T_c . The majority of the dissolved hydrogen forming the hydride precipitates comes from the acids used to etch the cavity surfaces when the bath temperature gets too high.

The typical cool down process used for the NPS cavity is to flow liquid nitrogen through the nitrogen shield as well as into the helium dewar. Once temperatures in the cavity approach 150 K, the onset of greatest hydrogen diffusion rate, we allow the liquid nitrogen to boil off and switch to putting liquid helium into the dewar. The goal is to pass through the temperature region of $66 \text{ K} \leq T \leq 150 \text{ K}$ as quickly as possible [12]. The result is a decreased risk of Q-disease at the expense of a significant boil off of helium. After five cycles between room temperature and 4 K, no decrease in cavity Q_0 has been observed to date.

Some of the experimental values that are typically desired from cold cavity testing are estimates of the cavity’s Q_0 value and the resulting cavity gradient. Plots of Q_0 values are also useful in diagnosing any issues that may be encountered during cold cavity testing, such as field emission, Q-disease, and multipacting.

a. Helium Processing

As part of the processing of the NPS cavity, the first ever helium processing in a commercial facility was performed. In helium processing, a small amount of helium gas is introduced into the cavity vacuum to serve as an ionization source [48], [49]. As electrons are emitted from field emission sites, the electrons ionize the helium gas causing the helium ions to back bombard the emission sites. The helium

ions have greater momentum and the resulting back bombardment of the emission sites increases the local temperature more than the ohmic heating and processes them down faster. Experience at Cornell [48] indicates they have observed factor of 2 increases in accelerating gradient using helium processing on their cavities. Estimates of accelerating gradient improvement in the NPS gun were just shy of this mark, improving the estimated accelerated gradient from 400 kV to 750 kV. The Q_0 plots calculated by Niowave compares this series of processing runs to those from the initial cold testing, as shown in Figure 106.

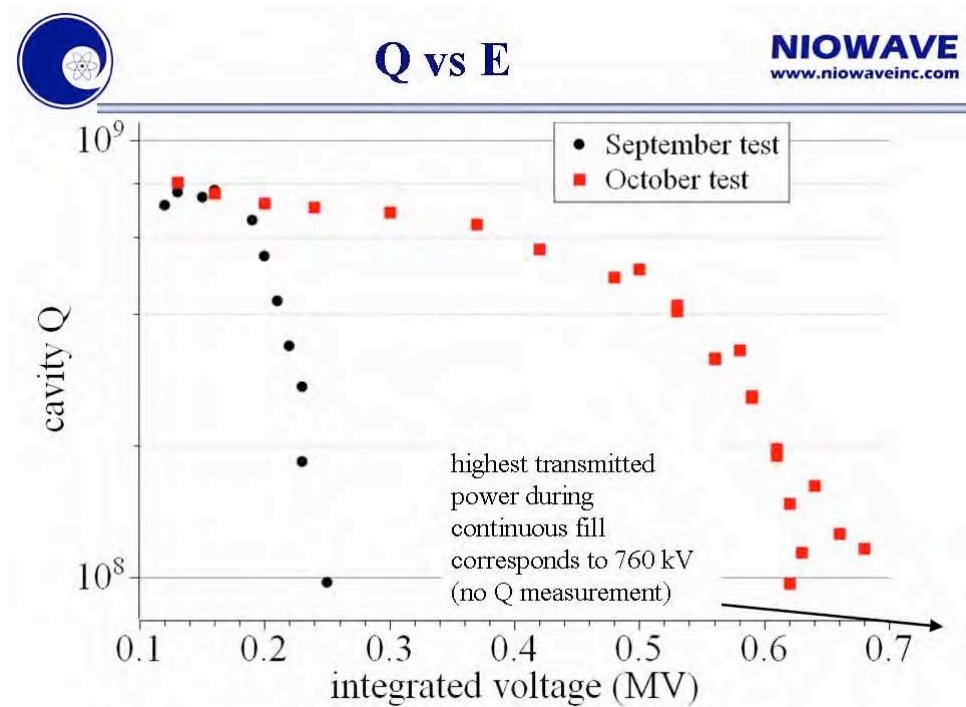


Figure 106. Two Q_0 vs. E plots for the NPS booster cavity are plotted. The black dots correspond to measurements made during the initial cold testing in September 2009. The red dots were taken after considerable RF and helium processing in October 2009. The September 2009 measurements were limited by X-ray generation. The October 2009 measurements were field emission limited.

5. Full Beam Test

After the cavity was vented and the cathode stalk installed as discussed in the previous chapter, all systems, including the drive laser, were ready for first beam testing. First beam testing was conducted at Niowave facilities the week of 07-11 June 2010 and established the NPS 500 MHz SRF gun as the first operational SRF gun in the United States. The gun transported an estimated 30 pC from a niobium cathode. Continued testing to fully characterize the gun has been ongoing since that date, but that data will not be reported here. Interested readers are encouraged to consult the literature for more recent published results.

The insertion of the cathode constitutes a major change to the cavity and could possibly affect how the system operated and behaved. Prior to conducting beam tests, we characterized the system and conducted processing runs to process any new field emission sites introduced as part of the venting process or present on the cathode assembly. While there were a few field emitters and multipacting barriers that required processing, the largest impact appeared to be the loss of cavity gradient due to coupling to the cathode stalk. With the cathode fully retracted (about 14 mm behind the nose cone), Q_0 measured approximately 5×10^8 . Adjustments to the cathode position placed it at an estimated retraction of about 9 mm behind the nose cone with a Q_0 of 6×10^7 . However, there is no way to currently describe the cathode position with confidence.

We know that changing the cathode position alters the fields inside the cavity and has an effect on the fundamental frequency of the cavity as well. Using the cathode itself as the probe, we can simulate the frequency shift as a function of cathode position over the range of desired positions, shown in Figure 107. Assuming we can verify this plot between beam performance characteristics and measured fundamental frequency against the simulations, we will have a useful diagnostic for determining

the position of the cathode. Full experimental verification has not yet been performed, however the first two data points indicate the slope of the curve appears to match the cavity frequency shift expected.

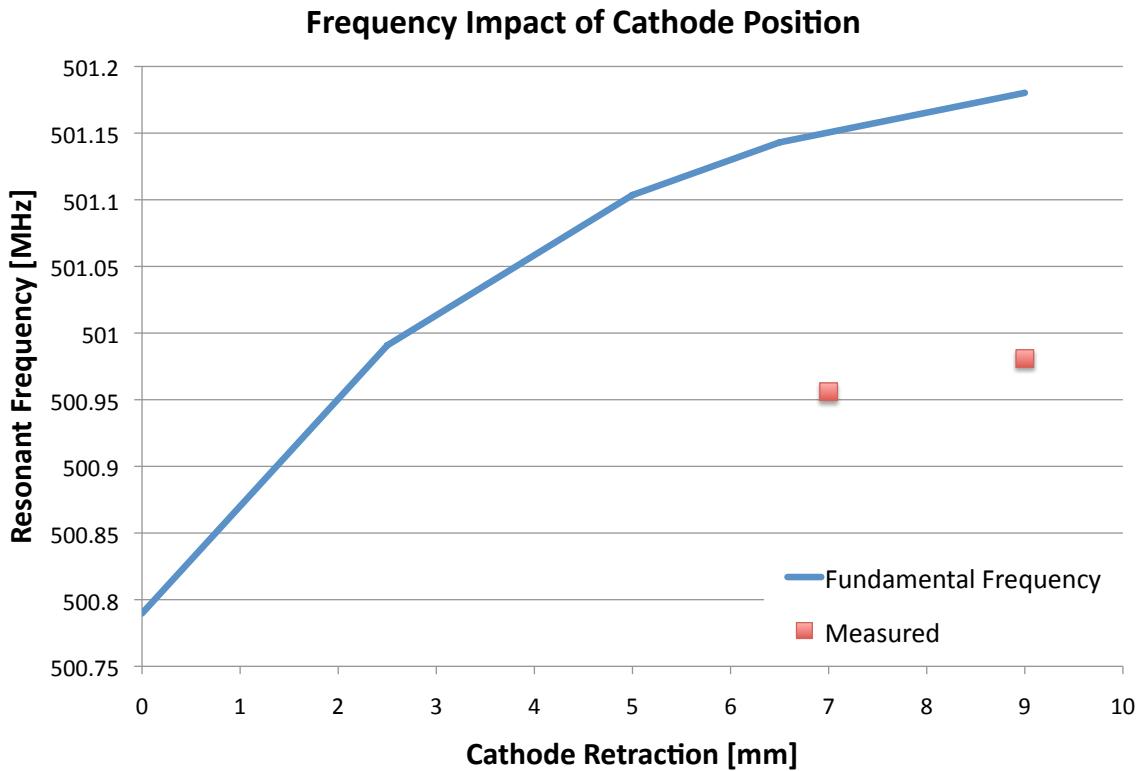


Figure 107. Results of Superfish simulations of the NPS SRF gun showing resonant frequency vs. cathode retraction. As the cathode position relative to the nose cone is adjusted, the fundamental frequency, f_0 , of the gun changes. It is possible that this effect could be used to measure cathode position once verified experimentally. The first two measured resonant frequencies as the cathode position is changed are plotted in red.

After RF processing, the gun was not able to return to field levels obtained in the booster configuration. It is believed that this is due primarily to the coupling of power out of the cavity by the cathode stalk. As the stalk was inserted nearer to the nose cone surface, the Q_0 of the cavity was seen to decrease, indicating a loss of stored power in the cavity. The cathode was returned to the 9 mm retracted

position and an estimated accelerating gradient of 200 kV reverified. At this point, it was decided that, although the final gap voltage was not as high as desired, it was enough for performing initial beam tests. First beam with the NPS 500 MHz SRF gun was achieved on the first attempt at 1110 EDT on 09 June 2010. Once beam was established, the team proceeded to commission the diagnostic beam line, verifying operation of all correctors and transporting electron beam to every instrumentation station in the beam line over the course of only 12 hours.

a. Instrumentation Stations

The instrumentation stations for the NPS gun beam line are the primary method for collecting data about the electron beam. There are five instrumentation stations along the main beam line and an additional two along the spectrometer leg. With the exception of the first station, the stations have identical construction. Figure 108 shows the pertinent portions of the stations. Only the first instrumentation station has the pinholes for performing emittance estimates.

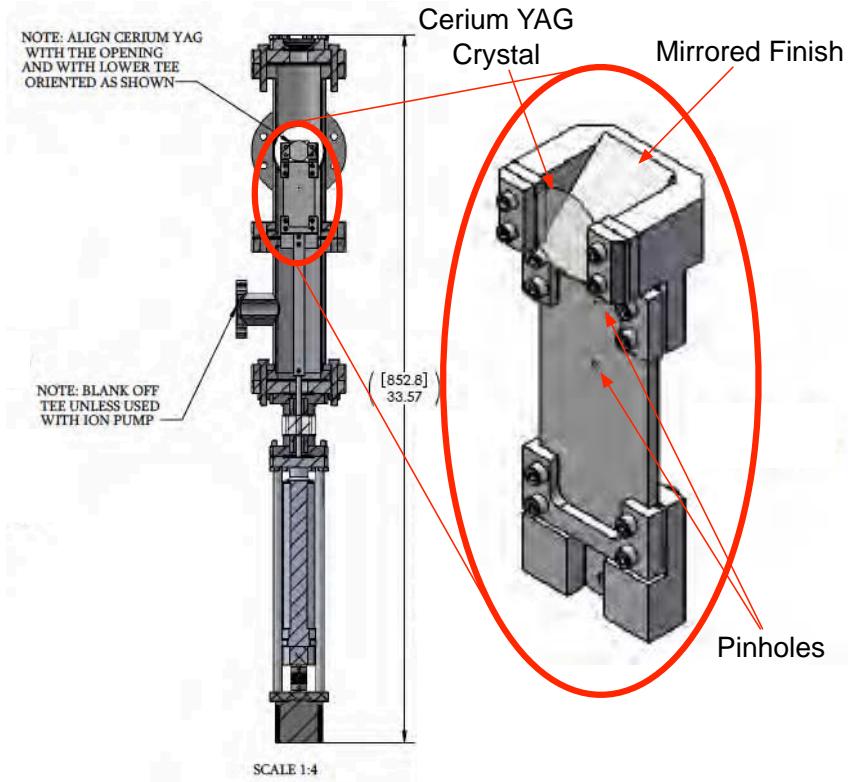


Figure 108. Each instrumentation station consists of an electrically isolated rod attached to a stepper motor driven linear actuator for insertion. Each station has a Cerium YAG crystal mounted in front of a polished aluminum mirror for beam viewing. The aluminum plate below the screen has pinholes on the first station for emittance measurements; for the others, this is a blank plate. The blank plate sections can be used as Faraday cups for measuring bunch charge.

Each station can be independently positioned to one of two instruments or retracted so as to present a clear aperture. The upper-most consists of a Cerium doped YAG screen which scintillates when struck by the electron beam. Cameras view the beam via top-mounted, down-looking video cameras via a polished aluminum mirror behind the Cerium YAG screen. Pin holes are installed in the first instrument station for performing emittance measurements. The bottom position of each station is intended for use a Faraday cup. Each station is electrically isolated, so any charge impacting the instrument carriage can be captured for bunch charge measurements.

Figure 109 shows some representative screens from the first beam test. For relative size measurements, the white line in each picture is the side view of the screen and each screen is 25.4 mm in length.

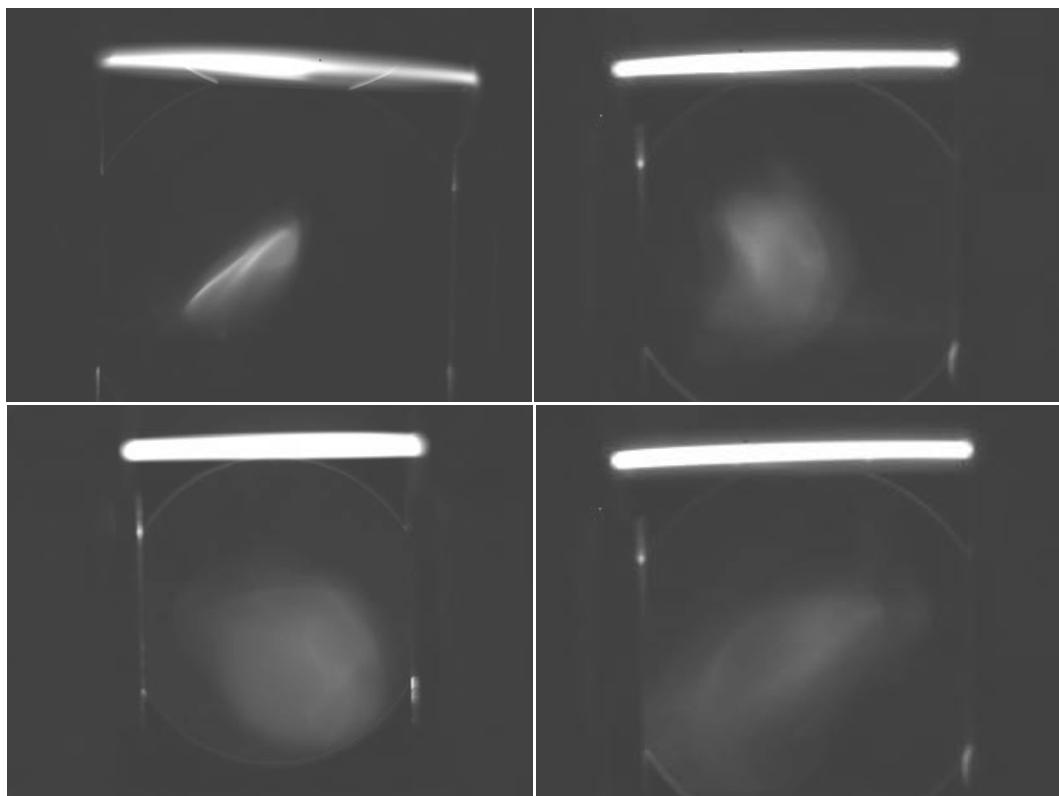


Figure 109. Beam cross sections captured from stations 2 (bottom left), 3 (top left), and 4 (right) during the first beam tests of the NPS 500 MHz SRF gun. Brightness is proportional to beam intensity. The bright line at the top of each picture is the side view of the Cerium YAG crystal.

b. Experimental Verification of Electron Beam Energy

To verify estimates of the cavity gradient, we prepared simulations to calculate the beam energy using the “window-frame” beam correctors. window-frame correctors are solenoid coils wrapped around a square iron ring, as seen in Figure 110. The solenoids are operated in pairs; each pair provides a deflecting field in one plane. While there are multiple window-frame correctors installed along the beam line, the intent is to use the first one to deflect the beam in the vertical plane (to avoid impinging on the laser mirror). Based on the same principle of the solenoid calculation, a beam of specific energy has a certain rigidity or resistance to deflection. For the same magnetic field a beam with a higher energy will deflect less than a similar beam with lower energy.

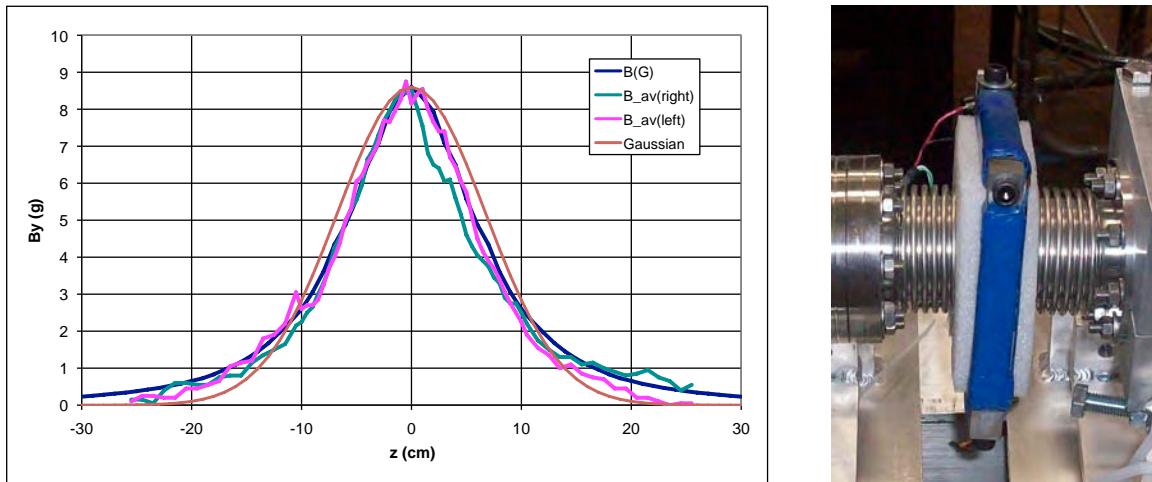


Figure 110. On the right is the window-frame corrector as installed on the diagnostic beam line. On the left is the field map through the axis of the corrector measured in the same manner as the solenoid field. A Gaussian fit to the field profile is plotted as well, for comparison.

Based on this concept, a series of simulated beams of multiple energies, using the NPS gun as the source, were propagated down a reduced beam line (terminating at the second diagnostic station). To model the window-frame corrector in the absence of actual field maps, a Gaussian approximation to the field was used.

This approximation was generated numerically in MATLAB, exported to a text file and imported into GPT. Actual field maps were inaccessible for these simulations as GPT and Superfish both require solenoidal fields to have their magnetic axis aligned to the beam propagation direction. For simplification and expedience, an approximate method was used. Hand calculations using an impulse approximation were performed to verify the simulation results. Figure 111 shows the results of the simulation runs, providing a simple process for determining beam energy based on the measured centroid deflection on the YAG screen and the applied corrector current.

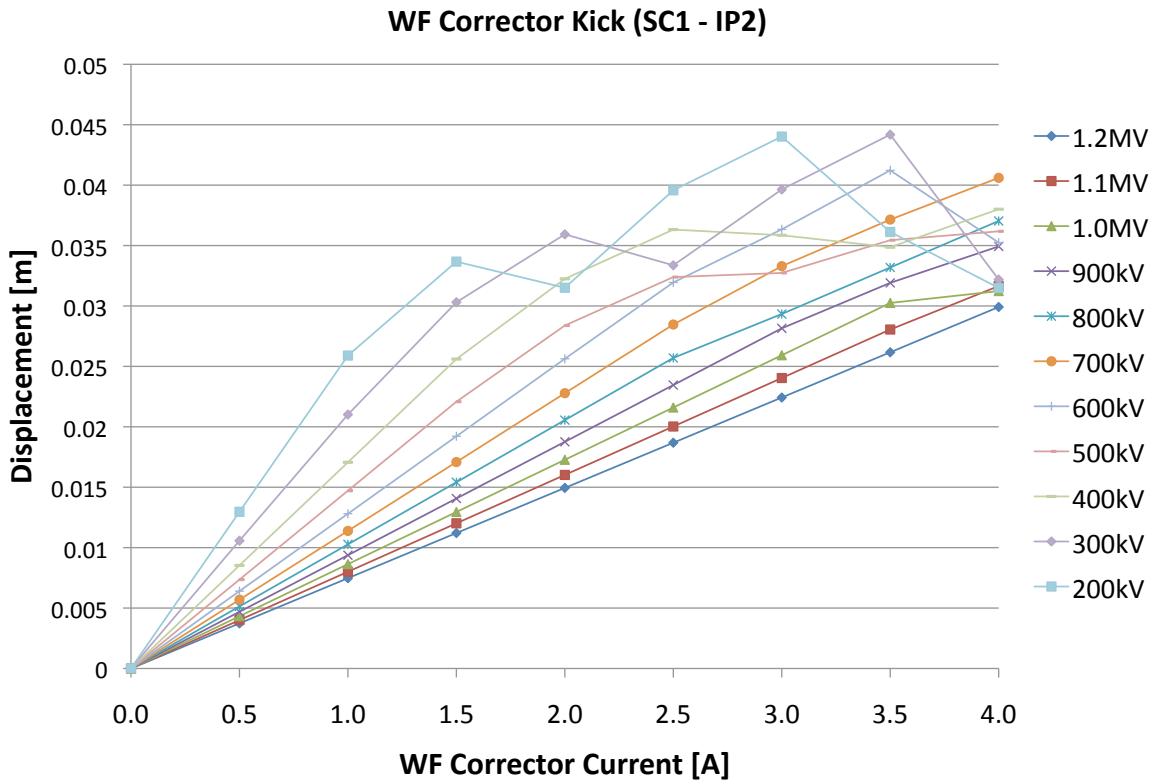


Figure 111. A look-up chart for determining beam energy based on corrector current and centroid deflection. The ragged values near 0.035 m deflection is due to beam clipping on the beam tube wall during propagation.

The first estimate of beam energy was obtained after commissioning the beam line to the first instrumentation station. With a corrector current of ~ 0.66

A, we observed a deflection of ~ 7.5 mm. Correcting this deflection for the additional propagation distance to the second instrumentation station and consulting the chart in Figure 111, we estimated the beam energy at 400 keV. The impulse calculation confirmed this result. However, RF calculations by Professor Brian Rusnak estimated the maximum beam energy as no higher than ~ 200 keV. Once the remainder of the beam line was commissioned, the dipole magnet was used to perform a second energy measurement, finding the beam energy to be approximately 200 keV.

Attempts to duplicate the measurement on the following day after performing further RF conditioning were complicated by difficulty transporting the beam to the termination of the spectrometer leg. Using a sweep of the beam through the first instrumentation station after the dipole, along the spectrometer leg, indicated a beam energy of approximately 250 keV.

Possible sources of error in the window-frame energy measurement could be misalignment of the beam prior to entering the corrector field. If the beam is already pointed off the beam line axis, any deflection measurement would be skewed by this error. Additionally, it was determined that the light-tight box around the UV drive laser insertion table could be magnetized, which could affect the accuracy of measurements performed with the window-frame correctors in this area of the beam line. However, with correlation between the dipole beam energy measurements and the RF calculations, we feel confident that the NPS SRF gun has developed a minimum 200 keV beam of electrons.

c. Experimental Verification of Bunch Charge

Calculation of bunch charge in the SRF beam line is performed in a similar fashion to that described in the DC gun beam line. In this case, there is only a single metal plate upon which the beam impinges. To verify that the beam is centered on the Faraday cup section, the electron beam is first steered and focused

onto the Cerium YAG screen. Then the Faraday cup is moved into the proper measurement position. Traces on the oscilloscope were almost rectangular and time broadened (about 10 ns), as seen in Figure 112.

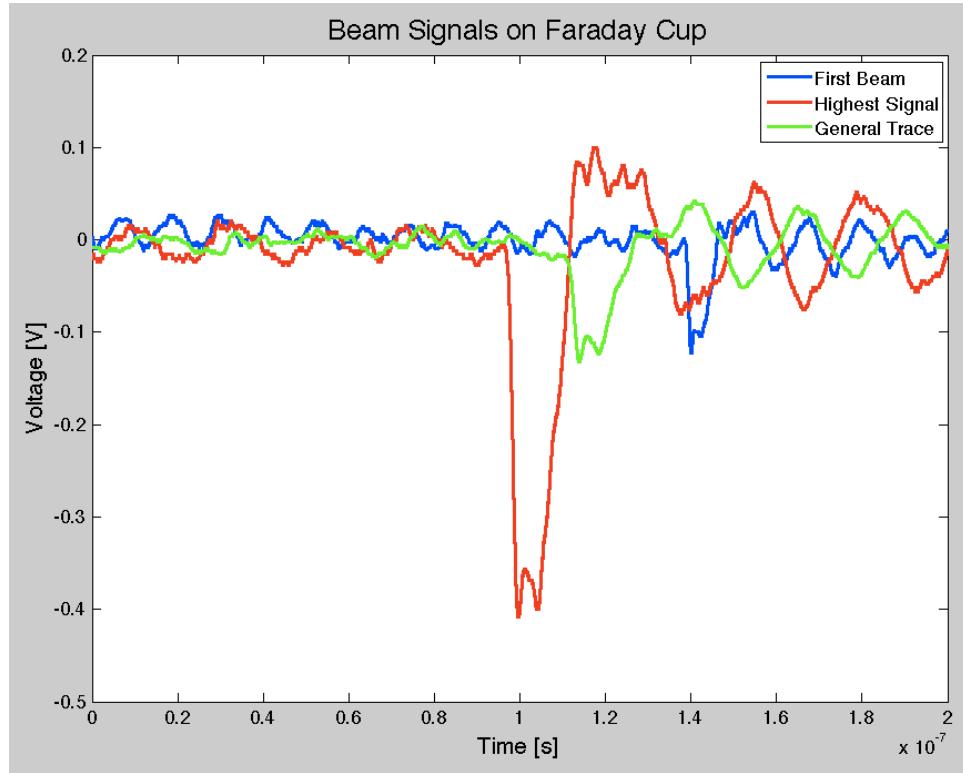


Figure 112. Faraday cup signal plots for three different cases. The blue is the signal from 1st beam. In red is the maximum signal strength from the same system set up after adjusting the laser position on the cathode. The green is a general plot from later in the experimental runs.

From a photodiode calibration, our Boeing collaborators were able to tell us the laser pulse energy for these three plots was about $180 \mu\text{J}$ per pulse. Using this value with the same process as shown in equation (IV.18), and modeling the pulse as a rectangular pulse, we find the bunch charges and quantum efficiency estimates shown in Table 7. Additionally, after adjustment of the laser and further transport optimization, the largest voltage pulse obtained during this series of runs is shown as well. The QE measurements are about an order of magnitude less than

anticipated when compared to values in the literature. However, the majority of the literature values are reported after having had some kind of processing of the surface beyond the BCP that was performed on ours. Future work includes performing laser cleaning/preparation of the cathode surface and these measurements provide us with a baseline value with which we can measure the improvement seen. It should also be noted that these values are preliminary in that the transport and launch phase are only roughly optimized. Any adjustment resulting in increased charge transport will show a corresponding increase in QE. It is possible that further adjustments will improve these measurement values prior to any cathode surface alterations.

Table 7. NPS gun bunch charge and QE measurements from the first experimental run.

Run Name	V [mV]	E_{pulse} [μ J]	q [pC]	QE
1st Beam	100	180	20	5.2×10^{-7}
Highest Signal	370	180	74	1.9×10^{-6}
General	110	180	22	5.7×10^{-7}
“Best”	550	300	110	1.7×10^{-6}

d. Bunch Charge vs. Launch Phase

Using the Faraday cup in the second instrumentation station as our detector, we can plot an acceptance plot of the gun for the 9 mm retracted cathode position. We adjust the launch phase by changing the phase delay between the RF signal and the laser pulse by adjusting a manual RF phase shifter. Using this process, we performed this measurement twice, once for an estimated 250 keV beam and once for a 300 keV beam with the plots shown in Figure 113. We see that the cavity has a relative phase acceptance of about 150° for the 250 keV beam and 165° for the 300 keV beam as measured between roll-off points. Comparing these acceptance window widths with simulation results reported earlier (of about 140°) we find them to be in good agreement.

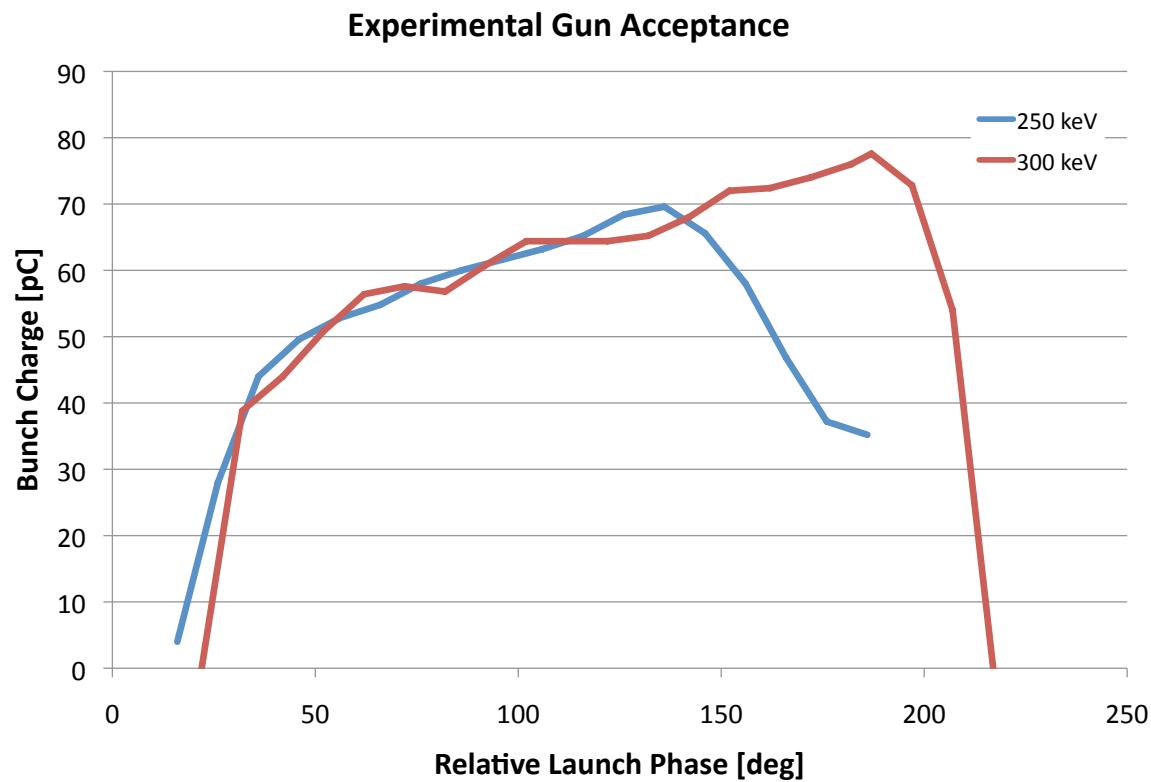


Figure 113. NPS gun relative phase acceptance for beam energies of 250 keV and 300 keV. Bunch charge was measured at the second instrumentation station.

THIS PAGE INTENTIONALLY LEFT BLANK

V. FEL BACKGROUND

In the next chapter, a scheme for describing the optical fields of a free electron laser (FEL) will be described. This chapter will provide a moderate background for the reader unfamiliar with FELs, but will not perform detailed derivations. For the interested reader, there are many accessible texts and papers [50], [51], [52] available for further study. This background chapter will start with the basic history of the FEL leading to a discussion of the major components of most FEL systems and the basic physics of the FEL.

A. HISTORY AND COMPONENTS

The FEL, as its name implies, is a laser. However, it is unlike other lasers in that the lasing medium does not have specific energy levels in which a population inversion must be established. In a typical laser, a pumping mechanism, whether a chemical reaction, flash lamps, or other technique is used to elevate electrons to an excited state. Once in this excited state, photons are collected and amplified through stimulated emission as the elevated energy level is depopulated.

1. History

The technology of the FEL has its origins in the microwave tubes of the 1930s, predating the solid state laser. With microwave tubes, scientists discovered they could generate coherent radiation from the oscillations of electrons. This coherent radiation, typically in the RF wavelengths was collected and passed into waveguides for many purposes. Depending on the size of the resonator cavities used, the microwave tube could be used to generate wavelengths of that fundamental size. This limited these tubes to wavelengths associated with resonant cavities that could be easily manufactured, typically on the order of centimeters. Although it was recognized that coherent radiation could be obtained from a beam of free electrons, the connection between

the microwave tube and the laser, first demonstrated in 1960, would take another 10 years before the concepts were combined.

The first laser, demonstrated at Hughes Research Labs [53] in 1960, was a solid state laser. The lasing medium was a solid laser crystal, pumped with a flash lamp. An optical resonator was put around the crystal to provide feedback during amplification. Gas lasers, chemical lasers, solid state lasers, and almost all other lasers operate in a similar fashion. Each of these lasers is dependent on the lasing medium properties, usually requiring shut down or reduced duty cycle to avoid over-heating the lasing medium or replenishment of the constituent chemicals. Additionally, every laser's operating wavelength is fixed to the lasing medium. If a different wavelength is desired, new constituents must be found.

As initially proposed [54] and demonstrated [55], an electron beam passing through a periodic magnetic field can experience stimulated emission, resulting in a laser that has no fixed lasing medium. The fact that the laser wavelength is established through an adjustable system, the electron accelerator, the FEL can conceivably be designed to operate at any desired wavelength. In fact, today FELs have operated at wavelengths from cm (microwaves) to 1.5 Å (hard X-ray) and nearly all wavelengths in between [56]. There have been many advances in technology since the first FEL lased at Stanford University in 1976.

Over the years of development, the FEL output power level has not received as much attention as other laser types. There probably is not a single primary reason other than the excitement of being able to design a laser for any particular part of the spectrum kept the investment money for FELs focused primarily on driving to shorter wavelengths. Since most conventional lasers share the same optical principles, once a laser was made to work at a particular wavelength, the main investment focus was in improving efficiency and output power. The FEL is finally receiving some

of the same attention, due primarily to military interest, driving FEL output power from watts, to kilowatts [57], to tens of kilowatts [58], with designs for a 100 kW laser underway [59].

2. FEL Components

All FELs have some common components that are important to the overall system performance. Some basic description of the important, major portions of a notional FEL will be discussed with the next section describing more of the physics and mathematical details. Most FELs can be described by the system layout similar to that shown in Figure 114 from Jefferson Laboratory, Newport News, VA.

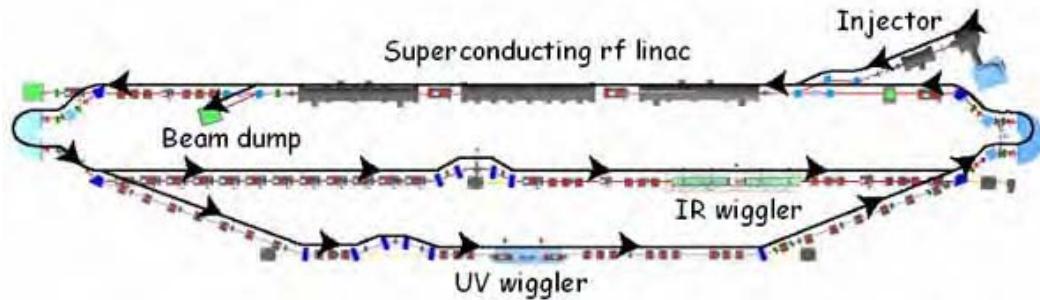


Figure 114. A system layout diagram of the JLAB FEL. While specific to one machine, this layout can be extrapolated to many other FEL systems. From [60].

a. *Injector*

All FELs use free electrons as their lasing medium. Typically the source of electrons is located in what is called the “gun.” As explained earlier, the two primary types of guns are radiofrequency (RF) and constant voltage (DC). Both of which can have different types of cathodes and can be normal, superconducting, or even a combination. There are trades-offs that must be carefully considered for both, but both types of guns are in use. Thermionic and photoelectric cathodes are the two primary electron sources in the guns and, again, there are trades necessary in determining the best type for the application intended.

The injector's primary purpose is to bring the energy of the bunch launched from the gun high enough to avoid slippage with respect to the RF fields as the bunch enters the primary linear accelerator (LINAC) sections. The injector may also have a dispersive element to adjust the electron bunch phase space for proper matching with the main LINAC.

b. Linear Accelerator

In the LINAC is where the electrons are given their primary boost to high energy. The LINAC can develop average kinetic energies from a few MeV to a many GeV depending on the wavelength desired in the end application. The only difference between machines is the amount of RF power put into the cavities and the operating frequency. The operating frequency will determine, to a certain extent, the maximum repetition rate at which the laser can operate—higher frequencies have more RF “buckets” for electron bunches. Currently, the standard rule of thumb is that the higher the repetition rate, the lower the bunch charge to prevent developing instabilities in the RF fields (wake fields) that would tend to break up the beam as more bunches pass through the LINAC. Multiple accelerating sections, or multiple passes through the same section, can be used to achieve the desired final beam energy.

c. Undulator

The undulator is the section of the FEL where the lasing interaction takes place. The undulator is a periodic magnetic field, which through the Lorentz force, causes the electrons to “wiggle.” This section of the FEL is also known as the “interaction region” and is directly analogous to the lasing medium of a standard laser. A diagram of a linear undulator is shown in Figure 115.

In all FELs, the undulator causes the electrons to radiate in the forward direction as they accelerate side-to-side. In an oscillator FEL, the spontaneous emission is collected at a mirror and reflected back to a second mirror which sends the light back into the undulator. The next electron bunch arrives at the same time

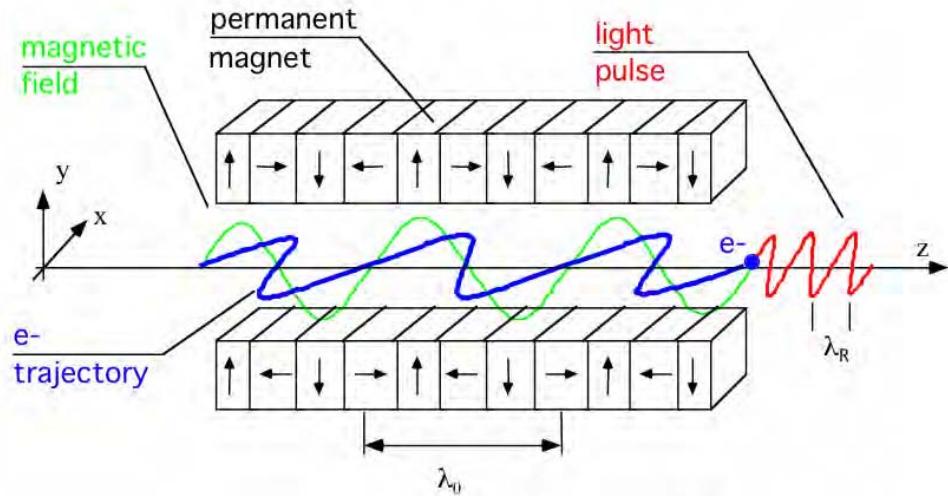


Figure 115. A linear undulator with alternating magnetic fields cause the electrons to oscillate in the opposite transverse plane. From [61].

as the light pulse and the electrons now oscillate in the presence of an electric field stimulating them to radiate at the same wavelength. Over many passes, this laser light builds up to significant levels.

SASE (Self-Amplified Spontaneous Emission) and amplifier FELs are single pass designs. In an amplifier FEL a seed laser of the desired output wavelength is chosen and sent through the undulator. The electrons see the electric field of the seed laser and amplify it through stimulated emission. In SASE, the electrons feed forward their spontaneous emission. As the spontaneous emission from the back of the electron bunch overtakes the front of the bunch, the front of the bunch amplifies this signal in the same manner described for the other types of FEL.

d. Beam Dump

The beam dump is essentially the completion of the electrical circuit that was started at the cathode. As the electron beam passed through the undulator and gave up energy to the laser, a significant energy spread is induced on the beam. Unfortunately, this means that the beam is unable to be used again for the FEL

interaction. In some systems, the beam is dumped into a metal block at the full remaining beam energy. This can result in extremely large amounts of radiation. Another option, as shown in Figure 114 and demonstrated in numerous FEL facilities, is to recirculate the electron beam through the LINAC 180° out of phase from the accelerating RF. Essentially, this recirculation scheme couples most of the left over energy of the electron beam back into the RF fields in the LINAC. Not only does this reduce the overall RF power required for the FEL, increasing overall efficiency, but it decreases the amount of kinetic energy dumped into the beam dump, which decreases the radiation generated.

B. BASIC FEL PHYSICS

In this section, we develop and examine the fundamental equations that govern FEL physics. We will start with the relativistic Lorentz force equations and apply them to an electron moving through the periodic magnetic field of the undulator. We will then discuss how an FEL is tunable by examining the resonance condition. Finally we will discuss the importance of peak current to the operation of an FEL.

1. Lorentz Equations

We start with the relativistic Lorentz equations, recognizing that for an FEL, we intend to use highly relativistic electrons,

$$\mathbf{F} = -e [\mathbf{E} + c(\boldsymbol{\beta} \times \mathbf{B})] = \frac{d\mathbf{p}}{dt} = \frac{d}{dt} (\gamma m \boldsymbol{\beta} c) , \quad (V.1)$$

where γ is defined in the usual manner,

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} . \quad (V.2)$$

2. Electron Motion in a Periodic Magnetic Field

Let us first consider an electron in a periodic magnetic field without an electric field. We also assume a helical undulator for simplicity in calculation; the same methodology can be used to solve for a linear undulator. We define our undulator

field as

$$\mathbf{B} = B \langle \cos(k_0 z), \sin(k_0 z), 0 \rangle , \quad (\text{V.3})$$

where $k_0 = 2\pi/\lambda_0$ and λ_0 is the wavelength of the undulator. Substituting equation (V.3) into equation (V.1), we find the equation for the acceleration experienced by the electron,

$$\dot{\beta} = \frac{eB}{\gamma m} \langle \beta_z \sin(k_0 z), -\beta_z \cos(k_0 z), -[\beta_x \sin(k_0 z) - \beta_y \cos(k_0 z)] \rangle . \quad (\text{V.4})$$

We look at the \hat{x} component of equation (V.4) and integrate to find

$$\begin{aligned} \int \dot{\beta}_x \, dt &= \int \frac{eB\beta_z}{\gamma m} \sin(k_0 z) \, dt \\ \beta_x &= \frac{-eB\lambda_0}{2\pi\gamma mc} \cos(k_0 z) + v_{x_0} . \end{aligned} \quad (\text{V.5})$$

If we assume perfect injection and helical motion, then $v_{x_0} = 0$. Otherwise, the electron would wander away from the axis at constant velocity. The β_y equation follows the same process with the same argument for the integration constant, giving

$$\beta_y = \frac{-K}{\gamma} \sin(k_0 z) , \quad (\text{V.6})$$

where

$$K \equiv \frac{eB\lambda_0}{2\pi mc} . \quad (\text{V.7})$$

Putting equations (V.5) and (V.6) into the \hat{z} component of equation (V.4), we find that for perfect injection,

$$\dot{\beta}_z = 0 , \quad (\text{V.8})$$

meaning that β_z is a constant, β_0 . We now know $z(t)$, and assuming that $t = 0$ corresponds to the start of the undulator and we call that position $z_0 = 0$, then $z(t) = c\beta_0 t$.

Let us now solve for the electron trajectories. We will solve the x component and state the y component when complete as they are solved in the same manner.

Starting with the $\dot{x} = c \beta_x$ equation,

$$\begin{aligned}\dot{x} &= \frac{-Kc}{\gamma} \cos(k_0 z) \\ x &= \frac{-K}{\gamma k_0 \beta_0} \sin(k_0 z) + x_0\end{aligned}\quad (\text{V.9})$$

and similarly,

$$y = \frac{K}{\gamma k_0 \beta_0} \cos(k_0 z) + y_0 . \quad (\text{V.10})$$

For perfect helical motion $x^2 + y^2 = r^2 = \text{a constant}$. This gives us

$$\begin{aligned}r^2 &= \left[\frac{-K}{\gamma k_0 \beta_0} \sin(k_0 z) + x_0 \right]^2 + \left[\frac{K}{\gamma k_0 \beta_0} \cos(k_0 z) + y_0 \right]^2 \\ r^2 &= x_0^2 + y_0^2 + \frac{2K}{\gamma k_0 \beta_0} [y_0 \cos(k_0 z) - x_0 \sin(k_0 z)] + \left(\frac{K}{\gamma k_0 \beta_0} \right)^2 ,\end{aligned}\quad (\text{V.11})$$

but this can only be a constant if the cross term in square brackets is equal to zero. The only way for this to be true is if $x_0 = y_0 = 0$. We now know the motion of the electrons to be

$$\mathbf{r}(t) = \left\langle \frac{-K}{\gamma k_0 \beta_0} \sin(k_0 \beta_0 c t), \frac{K}{\gamma k_0 \beta_0} \cos(k_0 \beta_0 c t), \beta_0 c t \right\rangle . \quad (\text{V.12})$$

3. Resonance Condition

One of the key characteristics of the FEL is its tunability. To demonstrate where that capability comes from, let us examine the resonance condition. The resonance condition is defined as being satisfied when one wavelength of light, λ , passes over an electron as the electron traverses one wavelength of the undulator, λ_0 . This can be thought of as a photon-electron race, described by the following relation,

$$\frac{\lambda_0}{\beta_z c} = \frac{\lambda + \lambda_0}{c}, \quad (\text{V.13})$$

which says that in the same amount of time it takes the electron to move one λ_0 , the photon will have traveled $\lambda + \lambda_0$. If we rearrange this equation, we find

$$\lambda = \lambda_0 \left(\frac{1 - \beta_z}{\beta_z} \right) . \quad (\text{V.14})$$

We recall the Lorentz factor and the same assumptions from the previous derivations for β_x and β_y to find

$$\begin{aligned}\beta^2 &= \beta_x^2 + \beta_y^2 + \beta_z^2 \\ &= \frac{K^2}{\gamma^2} [\sin^2(k_0 z) + \cos^2(k_0 z)] + \beta_z^2 \\ &= \frac{K^2}{\gamma^2} + \beta_z^2.\end{aligned}\tag{V.15}$$

We put this result into the definition of γ and solve for β_z ,

$$\begin{aligned}\gamma^2 &= \frac{1}{1 - \beta^2} \\ &= \frac{\gamma^2}{\gamma^2 - K^2 - \gamma^2 \beta_z^2} \\ \gamma^2 (1 - \beta_z^2) &= 1 + K^2 \\ \beta_z &= \sqrt{1 - \frac{1 + K^2}{\gamma^2}}.\end{aligned}\tag{V.16}$$

We now have β_z to put into equation (V.14). We see in equation (V.14) that β_z is in both the numerator and denominator. If we have a highly relativistic beam, then $\beta_z \simeq 1$ and $\gamma \gg 1$. We expand equation (V.16) in a Taylor series and substitute into equation (V.14) to find

$$\lambda = \lambda_0 \frac{1 - \beta_z}{\beta_z} \simeq \lambda_0 \left[1 - \left(1 - \frac{1 + K^2}{2\gamma^2} \right) \right] \simeq \frac{\lambda_0}{2\gamma^2} (1 + K^2).\tag{V.17}$$

From this relation, we see that to adjust the output light from the FEL, assuming the period of the undulator is fixed, we can either change the undulator parameter (move magnets closer/further away to increase/decrease B) or change the electron beam energy. In practice, individual FELs have demonstrated an ability to tune the output light over about an order of magnitude from their designed wavelength.

4. Current Effects

As the bunch charge passes through the undulator, it generates an optical wave due to the synchrotron radiation released as the electrons are accelerated transverse to their motion. The strength of the optical field is directly impacted by the amount

of current present in the beam. Without going through the derivation steps, we know that an electromagnetic wave equation, using the auxiliary field (and cgs units) can be written [50]

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial}{\partial t} \right) \mathbf{A} = -\frac{4\pi}{c} \mathbf{J}_\perp . \quad (\text{V.18})$$

This equation says that an electromagnetic wave is generated perpendicular to the direction of the motion of the charges in \mathbf{J} . This bodes well for us as our undulator provides the necessary motion to obtain radiation along the axis of propagation. It can be shown that the sum over the electron bunch which constitutes \mathbf{J} can be written as a dimensionless current, defined [50]

$$j \equiv \frac{8N(e\pi KL)^2 \rho}{\gamma^3 mc^2} , \quad (\text{V.19})$$

where N is the number of periods in the undulator, L is the total length of the undulator, and ρ is the electron density. The source term of equation (V.18) can then be written in dimensionless terms as

$$-\frac{4\pi}{c} \mathbf{J}_\perp \propto \langle -je^{-i\zeta} \rangle , \quad (\text{V.20})$$

where ζ is the phase of an electron relative to the optical field and $\langle \dots \rangle$ is an ensemble average over the electron phases.

From equation (V.19), we see that if we desire to drive the optical field using the bunch charge, we can increase the source term by increasing the bunch charge or maintaining the charge, but increasing the charge density within the bunch. We do this by focusing the electron beam to a tight spot radially and compressing the beam longitudinally just before the bunch enters the undulator. When one also takes into consideration the gain of the FEL and can get to the high gain regime ($j \gg \pi$), then the gain of the FEL becomes [50]

$$G(\tau) = \frac{e^{\tau\sqrt{3}} \sqrt[3]{\frac{j}{2}}}{9} , \quad (\text{V.21})$$

where τ is the dimensionless time defined ct/L . In the case of a high power FEL, as is desired for military applications, the interaction region will necessarily be required to

have high gain as out-coupling of the laser light will be significant. It is this concept that drives our desire for a high bunch charge, low emittance gun for use with the NPS FEL.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. HERMITE-GAUSSIAN DECOMPOSITION

Free electron lasers (FELs) generate electromagnetic radiation during the electron-wiggler interaction. This radiation then interacts with the electron beam to provide the feedback used in the FEL interaction. In oscillator FELs, the mirrors will interact with the optical mode which is fed back into the FEL interaction to be amplified. In an amplifier configuration, there is no feedback of radiation back into the interaction; instead, a seed laser or spontaneous emission provides an optical pulse and this radiation field is amplified.

Knowing the spatial mode composition of the radiation field is important in determining the propagation properties of the output optical field. There are numerous measures of beam quality which describe how well a beam propagates compared with a pure Gaussian, such as M^2 measurements and Strehl ratios (see [62] and [63] for examples). We desire a suitable basis set of normal modes with which to decompose the radiation field of interest from an FEL. Since oscillator FELs typically use a spherical mirror cavity, they lend themselves to a Hermite-Gaussian (HG) basis set. We desire to use them for the amplifier case as well. This application of the HG basis set to the amplifier is not without problems, and our procedure for overcoming these shortcomings in order to determine useful information will be discussed.

Through methods developed by Siegman [62], the HG modes are demonstrated as a suitable basis set of normal modes with which to decompose optical fields. Initial analysis of the more structured FEL oscillator case was explored and demonstrated by Vigil [64]. We further demonstrate the analysis is also applicable to the amplifier FEL and develop a “least modal composition” optimization scheme for fixing the two free basis scaling parameters. This theory is then coded and merged by the author with the NPS FEL4D simulation code and applied to simulations of current and proposed FEL experiments [1].

A. HG SOLUTIONS TO THE WAVE EQUATION

Let us begin from first principles and derive a wave equation from Maxwell's equations. We will then work through the reduced wave equation and into the paraxial case, which is applicable to FEL systems of interest. We will then show that the HG basis set is applicable to problems of this type and show how we can recast the wave equation for a preferred solution.

1. The Electromagnetic Wave Equation

Let us begin our wave equation search with Maxwell's equations (see [20] or [19]), stated below in MKS units,

$$\nabla \cdot \mathbf{D} = \rho, \quad (\text{VI.1})$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (\text{VI.2})$$

$$\nabla \cdot \mathbf{B} = 0, \quad (\text{VI.3})$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \quad (\text{VI.4})$$

where \mathbf{D} is the electric displacement field, ρ is the free charge density, \mathbf{E} is the electric field, \mathbf{B} is the magnetic field, \mathbf{H} is the auxiliary magnetic field, and \mathbf{J} is the free current density.

We assume that our fields of interest exist in vacuum, so we can use the constitutive relations $\mathbf{D} = \epsilon_0 \mathbf{E}$ and $\mathbf{B} = \mu_0 \mathbf{H}$ where ϵ_0 is the permittivity of free space and μ_0 is the permeability of free space. We can obtain a wave equation from either the electric or magnetic field equations. We choose the electric field for illustration and note that the derivation using the magnetic field follows in a similar manner.

Assuming \mathbf{E} and \mathbf{B} are point-wise continuous in space and continuous in time, we begin by taking the curl of equation (VI.2),

$$\begin{aligned}
\nabla \times (\nabla \times \mathbf{E}) &= \nabla \times \frac{-\partial \mathbf{B}}{\partial t} \\
\nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} &= \frac{-\partial}{\partial t} (\nabla \times \mathbf{B}) \\
\nabla \left(\frac{\rho}{\epsilon_0} \right) - \nabla^2 \mathbf{E} &= \frac{-\partial}{\partial t} \left(\mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \\
\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} &= \left(\mu_0 \frac{\partial \mathbf{J}}{\partial t} + \nabla \frac{\rho}{\epsilon_0} \right) , \tag{VI.5}
\end{aligned}$$

where in the last step, we used $\sqrt{\mu_0 \epsilon_0} = 1/c$. If we now confine our discussions only to areas outside the interaction region, then there are no free charges and no currents. Thus, the entire right hand side becomes 0, leaving us with the homogeneous partial differential equation,

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 . \tag{VI.6}$$

Now that we have a wave equation, we assume the form

$$\mathbf{E} (x, y, z, t) = \tilde{\mathbf{E}} (x, y) e^{i(kz - \omega t)} , \tag{VI.7}$$

where k is the wavenumber, and put this into the wave equation, we generate a much

more complicated differential equation. Calculating term by term, we find

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} = \frac{\partial^2 \tilde{\mathbf{E}}}{\partial x^2} e^{i(kz-\omega t)}, \quad (\text{VI.8})$$

$$\frac{\partial^2 \mathbf{E}}{\partial y^2} = \frac{\partial^2 \tilde{\mathbf{E}}}{\partial y^2} e^{i(kz-\omega t)}, \quad (\text{VI.9})$$

$$\begin{aligned} \frac{\partial^2 \mathbf{E}}{\partial z^2} &= \frac{\partial}{\partial z} \left[\frac{\partial \tilde{\mathbf{E}}}{\partial z} e^{i(kz-\omega t)} + ik \tilde{\mathbf{E}} e^{i(kz-\omega t)} \right] \\ &= \frac{\partial^2 \tilde{\mathbf{E}}}{\partial z^2} e^{i(kz-\omega t)} + ik \frac{\partial \tilde{\mathbf{E}}}{\partial z} e^{i(kz-\omega t)} + ik \left[\frac{\partial \tilde{\mathbf{E}}}{\partial z} e^{i(kz-\omega t)} + ik \tilde{\mathbf{E}} e^{i(kz-\omega t)} \right] \\ &= \left(\frac{\partial^2 \tilde{\mathbf{E}}}{\partial z^2} + 2ik \frac{\partial \tilde{\mathbf{E}}}{\partial z} - k^2 \tilde{\mathbf{E}} \right) e^{i(kz-\omega t)}, \end{aligned} \quad (\text{VI.10})$$

$$\begin{aligned} \frac{\partial^2 \mathbf{E}}{\partial t^2} &= \frac{\partial}{\partial t} \left[\frac{\partial \tilde{\mathbf{E}}}{\partial t} e^{i(kz-\omega t)} - i\omega \tilde{\mathbf{E}} e^{i(kz-\omega t)} \right] \\ &= \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} e^{i(kz-\omega t)} - i\omega \frac{\partial \tilde{\mathbf{E}}}{\partial t} e^{i(kz-\omega t)} - i\omega \left[\frac{\partial \tilde{\mathbf{E}}}{\partial t} e^{i(kz-\omega t)} - i\omega \tilde{\mathbf{E}} e^{i(kz-\omega t)} \right] \\ &= \left(\frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} - 2i\omega \frac{\partial \tilde{\mathbf{E}}}{\partial t} - \omega^2 \tilde{\mathbf{E}} \right) e^{i(kz-\omega t)}. \end{aligned} \quad (\text{VI.11})$$

Putting these into the wave equation from equation (VI.6) results in

$$\frac{\partial^2 \tilde{\mathbf{E}}}{\partial x^2} + \frac{\partial^2 \tilde{\mathbf{E}}}{\partial y^2} + \frac{\partial^2 \tilde{\mathbf{E}}}{\partial z^2} + 2ik \frac{\partial \tilde{\mathbf{E}}}{\partial z} - k^2 \tilde{\mathbf{E}} - \frac{1}{c^2} \left[\frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} - 2i\omega \frac{\partial \tilde{\mathbf{E}}}{\partial t} - \omega^2 \tilde{\mathbf{E}} \right] = 0. \quad (\text{VI.12})$$

If we now apply the slowly varying phase and amplitude approximation, assuming $\partial \tilde{\mathbf{E}}/\partial z$ and $\partial \tilde{\mathbf{E}}/\partial t$ are small on the scale of an optical wavelength, then we can say

$$\left| \frac{\partial^2 \tilde{\mathbf{E}}}{\partial z^2} \right| \ll k \left| \frac{\partial \tilde{\mathbf{E}}}{\partial z} \right| \quad \text{and} \quad \left| \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} \right| \ll \omega \left| \frac{\partial \tilde{\mathbf{E}}}{\partial t} \right|.$$

Applying this approximation and defining $\nabla_{\perp}^2 \equiv (\partial^2/\partial x^2 + \partial^2/\partial y^2)$ notation from before, we have a paraxial wave equation,

$$\left[\nabla_{\perp}^2 + 2ik \left(\frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} \right) \right] \tilde{\mathbf{E}} = 0. \quad (\text{VI.13})$$

We now change variables to a system that is commonly used in studying FEL interactions, altering the equation to “follow the light” by substituting $u = z - ct$. We also make dimensionless time by scaling τ to the time required for light to propagate through the undulator ($\tau \equiv ct/L$, where L is the length of the undulator). We now must change variables in the operator terms inside the parenthesis,

$$\begin{aligned}\frac{\partial}{\partial z} &= \frac{\partial u}{\partial z} \frac{\partial}{\partial u} + \frac{\partial \tau}{\partial z} \frac{\partial}{\partial \tau} = \frac{\partial}{\partial u} \\ \frac{\partial}{\partial t} &= \frac{\partial \tau}{\partial t} \frac{\partial}{\partial \tau} + \frac{\partial u}{\partial t} \frac{\partial}{\partial u} = \frac{c}{L} \frac{\partial}{\partial \tau} - c \frac{\partial}{\partial u} \\ \frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} &= \frac{\partial}{\partial u} + \frac{1}{c} \left(\frac{c}{L} \frac{\partial}{\partial \tau} - c \frac{\partial}{\partial u} \right) = \frac{1}{L} \frac{\partial}{\partial \tau}.\end{aligned}\quad (\text{VI.14})$$

Our wave equation now looks like

$$\left[\frac{-iL}{2k} \tilde{\nabla}_\perp^2 + \frac{\partial}{\partial \tau} \right] \tilde{\mathbf{E}} = 0, \quad (\text{VI.15})$$

which contains all longitudinal effects in the τ derivative term.

Since we have made dimensionless time, let us make the transverse spatial terms dimensionless as well. We choose the scale factor $k/2L$ so that

$$\begin{aligned}\tilde{x} &= x \sqrt{\frac{k}{2L}} \\ \tilde{y} &= y \sqrt{\frac{k}{2L}} \\ \tilde{\nabla}_\perp^2 &= \frac{\partial^2}{\partial \tilde{x}^2} + \frac{\partial^2}{\partial \tilde{y}^2} \\ \nabla_\perp^2 &= \frac{k}{2L} \tilde{\nabla}_\perp^2,\end{aligned}$$

and our dimensionless wave equation becomes

$$\left[\tilde{\nabla}_\perp^2 + 4i \frac{\partial}{\partial \tau} \right] \tilde{\mathbf{E}} = 0. \quad (\text{VI.16})$$

2. HG Solutions

Following the development of Siegman [62] and Vigil [64], we establish the HG polynomials as solutions to equation (VI.16). Using a general approach that will

be used for the higher-order modes later, we first show that the zeroth order HG polynomial (which is a “pure” Gaussian) is a solution. Let us define a dimensionless field $a(\tilde{\mathbf{r}}, \tau)$ such that

$$a(\tilde{\mathbf{r}}, \tau) = a_0 e^{-\left(p(\tau) + \frac{\tilde{r}^2}{\tilde{z}_0 q(\tau)}\right)}, \quad (\text{VI.17})$$

where $a(\tilde{x}, \tilde{y}, \tau) \propto \tilde{\mathbf{E}}(\tilde{x}, \tilde{y}, \tau)$, $\tilde{r}^2 = \tilde{x}^2 + \tilde{y}^2$, $\tilde{z}_0 = Z_0/L$, Z_0 is the Rayleigh length, $q(\tau)$ is a function describing the complex radius of curvature, and $p(\tau)$ is a function that describes a distance dependent scaling factor. We stipulate the initial condition that the field is Gaussian at some time $\tau = 0$, implying $p(0) = 0$ and $q(0) = 1$. Plugging into equation (VI.16), we find

$$\begin{aligned} \frac{\partial^2 a}{\partial \tilde{x}^2} &= a \left(\frac{-2}{\tilde{z}_0 q} + \frac{4\tilde{x}^2}{\tilde{z}_0^2 q^2} \right), \\ \frac{\partial^2 a}{\partial \tilde{y}^2} &= a \left(\frac{-2}{\tilde{z}_0 q} + \frac{4\tilde{y}^2}{\tilde{z}_0^2 q^2} \right), \\ \frac{\partial a}{\partial \tau} &= a_0 \left(\frac{-\partial p}{\partial \tau} e^{-p} e^{-\frac{\tilde{r}^2}{\tilde{z}_0 q}} + e^{-p} \left(\frac{\tilde{r}^2}{\tilde{z}_0 q^2} \frac{\partial q}{\partial \tau} \right) e^{-\frac{\tilde{r}^2}{\tilde{z}_0 q}} \right) \\ &= a \left(-\mathring{p} + \frac{\tilde{r}^2}{\tilde{z}_0 q^2} \mathring{q} \right), \text{ where } \mathring{\cdot} \text{ indicates } \partial/\partial\tau, \\ \left[\tilde{\nabla}_\perp^2 + 4i \frac{\partial}{\partial \tau} \right] a &= 0 \\ a \left(\frac{-2}{\tilde{z}_0 q} + \frac{4\tilde{x}^2}{\tilde{z}_0^2 q^2} \right) + a \left(\frac{-2}{\tilde{z}_0 q} + \frac{4\tilde{y}^2}{\tilde{z}_0^2 q^2} \right) &= -4ia \left(-\mathring{p} + \frac{\tilde{r}^2}{\tilde{z}_0 q^2} \mathring{q} \right) \\ \frac{-4}{\tilde{z}_0 q} + \frac{4(\tilde{x}^2 + \tilde{y}^2)}{\tilde{z}_0^2 q^2} &= 4i \left(\mathring{p} - \frac{\tilde{r}^2}{\tilde{z}_0 q^2} \mathring{q} \right) \\ \frac{\tilde{r}^2}{\tilde{z}_0^2 q^2} + \frac{i\tilde{r}^2}{\tilde{z}_0 q^2} \mathring{q} &= i \mathring{p} + \frac{1}{\tilde{z}_0 q} \\ \frac{\tilde{r}^2}{\tilde{z}_0 q^2} \left(\frac{1}{\tilde{z}_0} + i \mathring{q} \right) &= \left(i \mathring{p} + \frac{1}{\tilde{z}_0 q} \right). \end{aligned} \quad (\text{VI.18})$$

Since \tilde{r} can vary across the entire image plane at all times τ , then the only way this solution can be valid for all \tilde{r} and τ is if both sides equal 0 simultaneously. This provides us with a system of two equations for solving for p and q . Starting with the

q equation,

$$\begin{aligned}
\frac{1}{\tilde{z}_0} + i \dot{\tilde{q}} &= 0 \\
\int_{q_0}^q dq' &= \int_0^\tau \frac{i}{\tilde{z}_0} d\tau' \\
q &= q_0 + \frac{i\tau}{\tilde{z}_0} \\
q &= 1 + \frac{i\tau}{\tilde{z}_0}, \text{ since } q_0 = q(0) = 1.
\end{aligned} \tag{VI.19}$$

We use this result in our solution to the p equation,

$$\begin{aligned}
i \dot{\tilde{p}} + \frac{1}{\tilde{z}_0 q} &= 0 \\
\dot{\tilde{p}} &= \frac{i}{\tilde{z}_0 q} \\
\dot{\tilde{p}} &= \frac{i}{\tilde{z}_0 \left(1 + \frac{i\tau}{\tilde{z}_0}\right)} \\
\dot{\tilde{p}} &= \frac{i\tilde{z}_0 + \tau}{\tilde{z}_0^2 + \tau^2} \\
\int_{p_0}^p dp' &= \int_0^\tau \left(\frac{i\tilde{z}_0}{\tilde{z}_0^2 + \tau'^2} + \frac{\tau'}{\tilde{z}_0^2 + \tau'^2} \right) d\tau' \\
p &= i \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right) + \ln \sqrt{\frac{\tau^2 + \tilde{z}_0^2}{\tilde{z}_0^2}}, \text{ since } p_0 = p(0) = 0.
\end{aligned} \tag{VI.20}$$

It is apparent from the the fraction in the exponent of the field form (VI.17) and the solution (VI.19) that we can define a dimensionless mode radius $\varrho(\tau)$ such that

$$\varrho(\tau) = \sqrt{\tilde{z}_0 + \frac{\tau^2}{\tilde{z}_0}}. \tag{VI.21}$$

If we call $\varrho(0) = \varrho_0 = \sqrt{\tilde{z}_0}$, we can recast our equation for $q(\tau)$ in terms of ϱ ,

$$\frac{1}{q(\tau)} = \frac{\tilde{z}_0}{\tilde{z}_0 + \frac{\tau^2}{\tilde{z}_0}} - \frac{i\tau}{\tilde{z}_0 + \frac{\tau^2}{\tilde{z}_0}} = \frac{\varrho_0^2}{\varrho^2(\tau)} - \frac{i\tau}{\varrho^2(\tau)}. \tag{VI.22}$$

Making the appropriate substitutions into equation (VI.17), we find

$$\begin{aligned}
a(\tilde{\mathbf{r}}, \tau) &= a_0 \exp \left[-i \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right) - \ln \sqrt{\frac{\tau^2 + \tilde{z}_0^2}{\tilde{z}_0^2}} - \frac{\tilde{r}^2}{\tilde{z}_0 \left(1 + \frac{i\tau}{\tilde{z}_0} \right)} \right] \\
&= a_0 \left(\sqrt{\frac{\tau^2 + \tilde{z}_0^2}{\tilde{z}_0^2}} \right)^{-1} \exp \left[-i \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right) - \frac{\tilde{r}^2}{\tilde{z}_0} \left(\frac{\varrho_0^2}{\varrho^2} - \frac{i\tau}{\varrho^2} \right) \right] \\
&= a_0 \left(\frac{1}{\sqrt{\tilde{z}_0}} \sqrt{\tilde{z}_0 + \frac{\tau^2}{\tilde{z}_0}} \right)^{-1} \exp \left[\frac{-\tilde{r}^2}{\varrho^2} + i \left(\frac{\tau \tilde{r}^2}{\varrho_0^2 \varrho^2} - \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right) \right) \right] \\
&= a_0 \frac{\varrho_0}{\varrho} \exp \left[\frac{-\tilde{r}^2}{\varrho^2} + i \left(\frac{\tau \tilde{r}^2}{\varrho_0^2 \varrho^2} - \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right) \right) \right], \tag{VI.23}
\end{aligned}$$

which has the functional amplitude and phase form we desire. We can now define the phase

$$\phi(\tilde{\mathbf{r}}, \tau) \equiv \frac{\tilde{r}^2 \tau}{\varrho_0^2 \varrho^2} - \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right), \tag{VI.24}$$

so that

$$a(\tilde{\mathbf{r}}, \tau) = \underbrace{a_0 \frac{\varrho_0}{\varrho} e^{-\frac{\tilde{r}^2}{\varrho^2}}}_{\text{underbrace}} \overbrace{e^{i\phi(\tilde{\mathbf{r}}, \tau)}}^{\text{overbrace}}, \tag{VI.25}$$

where the underbrace indicates the Gaussian amplitude we expected from equation (VI.17) and the overbrace indicates the evolving phase term of the complex field.

3. Higher Order HG Modes

To account for higher order modes, we must adjust our general form for the dimensionless field, equation (VI.17). Let us introduce two new functions, h and g , which control the transverse behavior of the field in their respective arguments and define

$$a(\tilde{\mathbf{r}}, \tau) = a_0 g \left(\frac{\tilde{x}}{\varrho} \right) h \left(\frac{\tilde{y}}{\varrho} \right) e^{-\left(p + \frac{\tilde{r}^2}{\tilde{z}_0^2} \right)}. \tag{VI.26}$$

We apply the dimensionless wave equation (VI.16) to this more general field, but take a useful detour to go through the appropriate partial derivatives,

$$\begin{aligned}
\frac{\partial g}{\partial \tilde{x}} &= \frac{\partial g}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} \frac{\partial}{\partial \tilde{x}} \left(\frac{\tilde{x}}{\varrho}\right) \\
&= \frac{1}{\varrho} \frac{\partial g}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} \\
&= \frac{1}{\varrho} g' \text{ where we have defined } g' \equiv \frac{\partial g}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} , \tag{VI.27}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 g}{\partial \tilde{x}^2} &= \frac{\partial}{\partial \tilde{x}} \left(\frac{\partial g}{\partial \tilde{x}} \right) = \frac{1}{\varrho} \frac{\partial}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} \left(\frac{1}{\varrho} \frac{\partial g}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} \right) \\
&= \frac{1}{\varrho^2} g'' , \\
\frac{\partial g}{\partial \tau} &= \frac{\partial g}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} \frac{\partial}{\partial \tau} \left(\frac{\tilde{x}}{\varrho}\right) \\
&= \tilde{x} \frac{\partial g}{\partial \left(\frac{\tilde{x}}{\varrho}\right)} \left(\frac{-1}{\varrho^2} \right) \frac{\partial \varrho}{\partial \tau} \\
&= \frac{-\tilde{x}}{\varrho^2} g' \overset{\circ}{\varrho} , \tag{VI.28}
\end{aligned}$$

and the h terms are derived in a similar fashion

$$\text{with } h' \equiv \frac{\partial h}{\partial \left(\frac{\tilde{y}}{\varrho}\right)} \text{ so that,}$$

$$\frac{\partial^2 h}{\partial \tilde{y}^2} = \frac{1}{\varrho^2} h'' \tag{VI.29}$$

$$\frac{\partial h}{\partial \tau} = \frac{-\tilde{y}}{\varrho^2} h' \overset{\circ}{\varrho} . \tag{VI.30}$$

Putting these terms into the dimensionless wave equation (VI.16) and performing some algebra,

$$\begin{aligned}
\frac{1}{\varrho^2} \frac{g''}{g} - \frac{4\tilde{x}}{\tilde{z}_0 q \varrho} \frac{g'}{g} - \frac{2}{z_0 q} + \frac{4\tilde{x}^2}{\tilde{z}_0^2 q^2} + \frac{1}{\varrho^2} \frac{h''}{h} - \frac{4\tilde{y}}{\tilde{z}_0 q \varrho} \frac{h'}{h} - \frac{2}{\tilde{z}_0 q} + \\
\frac{4\tilde{y}^2}{\tilde{z}_0^2 q^2} - \frac{4i\tilde{x}}{\varrho^2} \frac{g'}{g} \overset{\circ}{\varrho} - \frac{4i\tilde{y}}{\varrho^2} \frac{h'}{h} \overset{\circ}{\varrho} - 4i \overset{\circ}{p} + \frac{4i\tilde{r}^2}{\tilde{z}_0 q^2} \overset{\circ}{q} = 0 . \tag{VI.31}
\end{aligned}$$

We now rearrange these terms into something a bit easier to analyze,

$$\begin{aligned}
\left[\tilde{\nabla}_\perp^2 + 4i \frac{\partial}{\partial \tau} \right] a = 0 &= \underbrace{\frac{1}{\varrho^2} \frac{g''}{g} - \frac{4\tilde{x}}{\tilde{z}_0 q \varrho} \frac{g'}{g} - 4i\tilde{x} \frac{g'}{g} \frac{\dot{\varrho}}{\varrho^2}}_{\text{only } \tilde{x} \text{ dependence } (= -\alpha)} \\
&+ \underbrace{\frac{1}{\varrho^2} \frac{h''}{h} - \frac{4\tilde{y}}{\tilde{z}_0 q \varrho} \frac{h'}{h} - 4i\tilde{y} \frac{h'}{h} \frac{\dot{\varrho}}{\varrho^2}}_{\text{only } \tilde{y} \text{ dependence } (= -\beta)} \\
&+ \underbrace{\tilde{r}^2 \left(\frac{4i}{\tilde{z}_0 q^2} \dot{\tilde{q}} + \frac{4}{\tilde{z}_0^2 q^2} \right)}_{\text{only } \tilde{r} \text{ dependence } (=0)} \\
&- \underbrace{\left(\frac{4}{\tilde{z}_0 q} + 4i \dot{\tilde{p}} \right)}_{\text{no } \tilde{r} \text{ dependence}}. \tag{VI.32}
\end{aligned}$$

To have a solution to equation (VI.32) at some time τ , the coefficient of \tilde{r}^2 should be zero since it describes the field in the plane, and the remaining terms (“only \tilde{x} ”, “only \tilde{y} ”, and “no \tilde{r} dependence”) should sum to zero. We will let the \tilde{x} terms equal a constant, $-\alpha$, and the \tilde{y} terms equal a different constant, $-\beta$. We now solve these four differential equations starting with the terms proportional to \tilde{r} .

$$\begin{aligned}
\left(\frac{4i}{\tilde{z}_0 q^2} \dot{\tilde{q}} + \frac{4}{\tilde{z}_0^2 q^2} \right) &= 0 \\
\left(i \dot{\tilde{q}} + \frac{1}{\tilde{z}_0} \right) &= 0,
\end{aligned}$$

after factoring out and dividing both sides by

$$\frac{4}{\tilde{z}_0 q^2}.$$

This is the same equation we solved for q in the zeroth order solution, so it has the same solution based on the same conditions,

$$q(\tau) = 1 + \frac{i\tau}{\tilde{z}_0}. \tag{VI.33}$$

Next, let us focus on the \tilde{x} terms,

$$\begin{aligned}
\frac{1}{\varrho^2} \frac{g''}{g} - \frac{4\tilde{x}}{\tilde{z}_0 q \varrho} \frac{g'}{g} - 4i\tilde{x} \frac{g'}{g} \frac{\dot{\varrho}}{\varrho^2} &= -\alpha \\
g'' - \frac{4\tilde{x}\varrho}{\tilde{z}_0 q} g' - 4i\tilde{x} \frac{\dot{\varrho}}{\varrho} g' &= -\alpha g \varrho^2 \\
g'' - 4\tilde{x}g' \left(\frac{\varrho}{\tilde{z}_0 q} + i \frac{\dot{\varrho}}{\varrho} \right) + \alpha g \varrho^2 &= 0 \\
g'' - \frac{4\tilde{x}}{\varrho} g' + \alpha g \varrho^2 &= 0
\end{aligned} \tag{VI.34}$$

after substituting for q and $\dot{\varrho}$.

We recognize equation (VI.34) as Hermite's equation[65]. To recast equation (VI.34) in the explicit form of Hermite's equation, we introduce

$$\chi \equiv \frac{\tilde{x}\sqrt{2}}{\varrho} ,$$

so that

$$g' = \sqrt{2} \frac{\partial g}{\partial \chi} ,$$

and

$$g'' = 2 \frac{\partial^2 g}{\partial \chi^2} .$$

After substituting into (VI.34), we have the form we desire,

$$\frac{\partial^2 g}{\partial \chi^2} - 2\chi \frac{\partial g}{\partial \chi} + \frac{\alpha \varrho^2}{2} g = 0 ,$$

whose solutions are the Hermite polynomials with argument χ :

$$g(\chi) = H_m(\chi) = H_m \left(\frac{\tilde{x}\sqrt{2}}{\varrho} \right) , \tag{VI.35}$$

where $m \in \{\mathbb{J} \geq 0\}$ and \mathbb{J} represents the set of integers.

Similarly, the \tilde{y} terms from equation (VI.32) are solved to find

$$\begin{aligned}
\frac{1}{\varrho^2} \frac{h''}{h} - \frac{4\tilde{y}}{\tilde{z}_0 q \varrho} \frac{h'}{h} - 4i\tilde{y} \frac{h'}{h} \frac{\dot{\varrho}}{\varrho^2} &= -\beta \\
h'' - \frac{4\tilde{y}\varrho}{\tilde{z}_0 q} h' - 4i\tilde{y} \frac{\dot{\varrho}}{\varrho} h' &= -\beta g \varrho^2 \\
h'' - 4\tilde{y}h' \left(\frac{\varrho}{\tilde{z}_0 q} + i \frac{\dot{\varrho}}{\varrho} \right) + \beta h \varrho^2 &= 0 \\
h'' - \frac{4\tilde{y}}{\varrho} h' + \beta h \varrho^2 &= 0 \\
\frac{\partial^2 h}{\partial \chi^2} - 2\chi \frac{\partial h}{\partial \chi} + \frac{\beta \varrho^2}{2} h &= 0 \\
h(\chi) = H_n(\chi) &= H_n \left(\frac{\tilde{y}\sqrt{2}}{\varrho} \right), \tag{VI.36}
\end{aligned}$$

where we used $\chi \equiv \tilde{x}\sqrt{2}/\varrho$, $h' = \sqrt{2} \partial h / \partial \chi$, $h'' = 2 \partial^2 h / \partial \chi^2$, and $n \in \{\mathbb{J} \geq 0\}$. Additionally, for Hermite solutions, α and β from equations (VI.32), (VI.35) and (VI.36) are constrained by the Hermite equation form to

$$\frac{\alpha \varrho^2}{2} = 2m \text{ and } \frac{\beta \varrho^2}{2} = 2n. \tag{VI.37}$$

Moving to the last term in equation (VI.32),

$$\begin{aligned}
-\left(\frac{4}{\tilde{z}_0 q} + 4i \frac{\dot{p}}{p} \right) &= \alpha + \beta \\
\frac{4}{\tilde{z}_0 q} + 4i \frac{\dot{p}}{p} &= \frac{-4}{\varrho^2} (n + m) \\
i \frac{\dot{p}}{p} &= \frac{-1}{\tilde{z}_0 q} - \frac{1}{\varrho^2} (n + m) \\
\dot{p} &= i \left(\frac{1}{\tilde{z}_0 q} + \frac{1}{\varrho^2} (n + m) \right) \\
\dot{p} &= \frac{i}{\varrho^2} (m + n) + \frac{i}{\tilde{z}_0 + \frac{\tau^2}{\tilde{z}_0}} + \frac{\tau}{\tilde{z}_0^2 + \tau^2} \\
\int_{p_0}^p dp' &= \int_0^\tau \left(\frac{i\tilde{z}_0}{\tilde{z}_0^2 + \tau'^2} (m + n + 1) + \frac{\tau'}{\tilde{z}_0^2 + \tau'^2} \right) d\tau' \\
p &= i(m + n + 1) \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right) + \ln \sqrt{\frac{\tilde{z}_0^2 + \tau^2}{\tilde{z}_0^2}}. \tag{VI.38}
\end{aligned}$$

We finally have the pieces we need to state the HG solution to our dimensionless wave equation that includes higher-order modes. Putting everything together, we have

$$a(\tilde{\mathbf{r}}, \tau) = a_0 \frac{\varrho_0}{\varrho} H_m \left(\frac{\tilde{x}\sqrt{2}}{\varrho} \right) H_n \left(\frac{\tilde{y}\sqrt{2}}{\varrho} \right) e^{\frac{-\tilde{r}^2}{\varrho^2}} e^{i\phi(\tilde{\mathbf{r}}, \tau)}, \quad (\text{VI.39})$$

$$\phi(\tilde{\mathbf{r}}, \tau) = \frac{\tilde{r}^2 \tau}{\varrho_0^2 \varrho^2} - (m + n + 1) \tan^{-1} \left(\frac{\tau}{\tilde{z}_0} \right), \quad (\text{VI.40})$$

where the Hermite polynomials through order $m = 7$ are shown in Table 8 for reference.

Table 8. Hermite polynomials, H_n , through 7th order, where n indicates the order.

$H_0(x)$	$=$	1
$H_1(x)$	$=$	$2x$
$H_2(x)$	$=$	$4x^2 - 2$
$H_3(x)$	$=$	$8x^3 - 12x$
$H_4(x)$	$=$	$16x^4 - 48x^2 + 12$
$H_5(x)$	$=$	$32x^5 - 160x^3 + 120x$
$H_6(x)$	$=$	$64x^6 - 480x^4 + 720x^2 - 120$
$H_7(x)$	$=$	$128x^7 - 1344x^5 + 3360x^3 - 1680x$

At this point, we have shown the general and complete solution to the wave equation. We are almost ready to begin applying it to useful problems. However, it would be nice if we could adjust the form so that it had a more intuitive feel. Define $\Lambda \equiv \pi \varrho^2$ and $\Omega \equiv \tau / \tilde{z}_0$. These definitions will allow us to recast our solution in terms of two parameters, the mode area Λ and the mode phase Ω , for which we have an intuitive feel for how they should affect the field's behavior. After these substitutions, our solutions to the wave equation become,

$$a(\tilde{\mathbf{r}}, \Omega) = \frac{a_0}{\sqrt{1 + \Omega^2}} H_m \left(\tilde{x} \sqrt{\frac{2\pi}{\Lambda}} \right) H_n \left(\tilde{y} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{\frac{-\pi\tilde{r}^2}{\Lambda}} e^{i\phi} \quad (\text{VI.41})$$

$$\phi(\tilde{\mathbf{r}}, \Omega) = \frac{\pi\tilde{r}^2\Omega}{\Lambda} - (m + n + 1) \tan^{-1} \Omega. \quad (\text{VI.42})$$

The HG modes defined by the above equations form a complete, orthogonal basis [65], from which we can decompose and generate any optical field. We will normalize this equation in the next section. Figure 116 shows intensity patterns for a number of HG modes. Of note is that the order of the mode indicates the number of nulls in that coordinate direction.

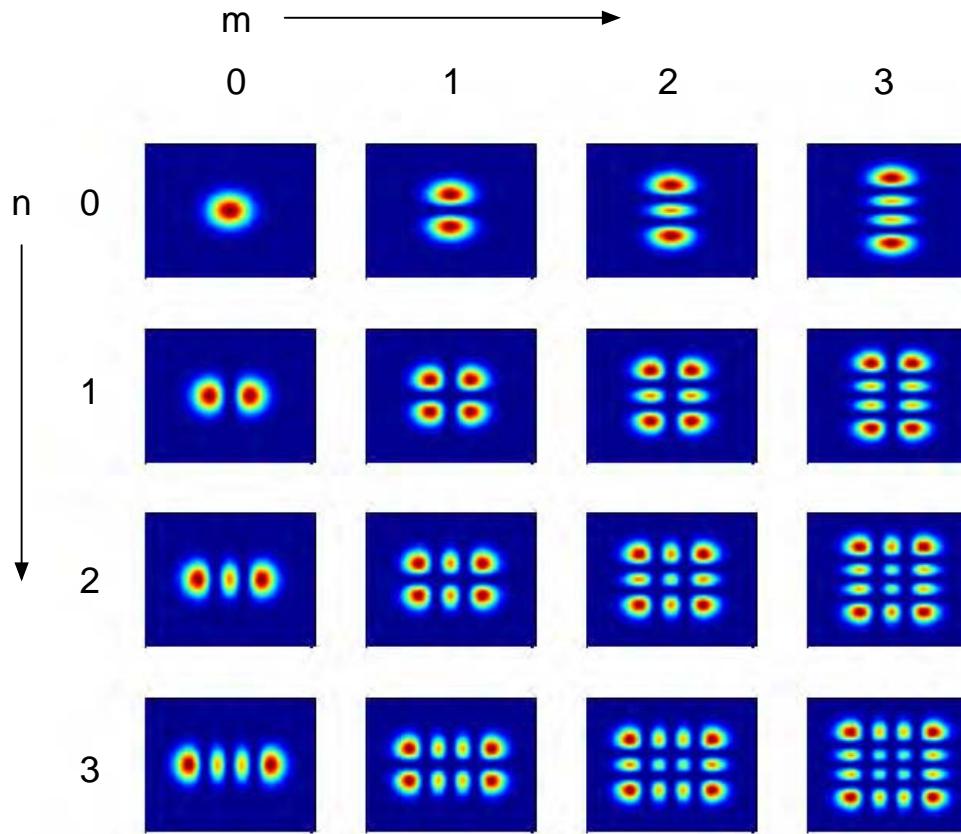


Figure 116. Intensity patterns for all HG modes through third-order in each coordinate. The top left image is the 0,0 mode with the mode number increasing to the left and down. The bottom right image is the 3,3 mode.

B. FIELD DECOMPOSITION

An arbitrary field can be decomposed with the HG basis set that we have shown to be a solution to the appropriate wave equation. This is done using an overlap integral to determine the complex coefficients, representing the “amount” of the (m, n) HG mode present in the field. However, for the decomposition to be deterministic, the Λ and Ω scaling factors must be fixed. Since Λ and Ω are related to p and q in our derivation, we have only recast them in more intuitive terms. We now address how to apply this decomposition method to representative fields and describe the process of fixing the scaling factors such that the minimum number of modes are required to describe an optical field of interest.

Since the overall amplitude of the field is not of importance for the decomposition, we scale the field such that the integrated intensity (aa^*) equals unity. This normalized field must then be projected against a set of normalized basis functions. In the previous section, we determined our basis functions, but we have not yet normalized them. To normalize equation (VI.41), we continue following the Vigil process [64] and start by determining the wavefront energy or power in the fundamental $(0,0)$ mode,

$$P_{0,0} = \iint_{-\infty}^{\infty} a_{0,0}^* a_{0,0} \, d\tilde{x} \, d\tilde{y} . \quad (\text{VI.43})$$

Putting equation (VI.41) into the integral, and substituting $H_0 = 1$,

$$\begin{aligned} P_{0,0} &= \frac{a_0^2}{1 + \Omega^2} \left[\int_{-\infty}^{\infty} H_0^2 \left(\tilde{x} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{-\frac{2\pi\tilde{x}^2}{\Lambda}} \, d\tilde{x} \right] \left[\int_{-\infty}^{\infty} H_0^2 \left(\tilde{y} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{-\frac{2\pi\tilde{y}^2}{\Lambda}} \, d\tilde{y} \right] \\ &= \frac{a_0^2}{1 + \Omega^2} \left(\int_{-\infty}^{\infty} e^{-\frac{2\pi\tilde{x}^2}{\Lambda}} \, d\tilde{x} \right) \left(\int_{-\infty}^{\infty} e^{-\frac{2\pi\tilde{y}^2}{\Lambda}} \, d\tilde{y} \right) , \end{aligned} \quad (\text{VI.44})$$

where we have made use of $H_0(x) = 1$. Recalling that

$$\int_{-\infty}^{\infty} e^{-bx^2} \, dx = \sqrt{\frac{\pi}{b}} , \quad (\text{VI.45})$$

we continue, setting $b = 2\pi/\Lambda$,

$$\begin{aligned}
P_{0,0} &= \frac{a_0^2}{1 + \Omega^2} \left(\sqrt{\frac{\pi}{b}} \right) \left(\sqrt{\frac{\pi}{b}} \right) \\
&= \frac{a_0^2 \pi}{b(1 + \Omega^2)} \\
&= \frac{a_0^2 \Lambda}{2(1 + \Omega^2)}.
\end{aligned} \tag{VI.46}$$

This is sufficient for the zeroth order HG polynomial, but does not address the higher-order modes. We now turn our attention there and recall that the HG polynomials comprise an orthonormal basis set with respect to a certain weighting function [65],

$$\int_{-\infty}^{\infty} H_i(x) H_j(x) e^{-x^2} dx = \delta_{i,j} 2^j j! \sqrt{\pi}, \tag{VI.47}$$

where $\delta_{i,j}$ is the Kronecker delta function.

Proceeding for $P_{m,n}$ as we did for $P_{0,0}$,

$$\begin{aligned}
P_{m,n} &= \iint_{-\infty}^{\infty} a_{mn}^* a_{mn} d\tilde{x} d\tilde{y} \\
&= \frac{a_0^2}{1 + \Omega^2} \left[\int_{-\infty}^{\infty} H_m^2 \left(\tilde{x} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{-\frac{2\pi\tilde{x}^2}{\Lambda}} d\tilde{x} \right] \times \\
&\quad \left[\int_{-\infty}^{\infty} H_n^2 \left(\tilde{y} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{-\frac{2\pi\tilde{y}^2}{\Lambda}} d\tilde{y} \right]
\end{aligned} \tag{VI.48}$$

where we note that the integrals in equation (VI.48) are of the same form as equation (VI.47) with the exception of the coefficients. If we define $\chi \equiv \tilde{x} \sqrt{2\pi/\Lambda}$, then $d\tilde{x} = \sqrt{\Lambda/2\pi} d\chi$.

Looking again at the first integral in equation (VI.48), and making note of the orthogonality and normalization relation of the HG polynomials, equation (VI.47), we solve

$$\begin{aligned}
I &= \int_{-\infty}^{\infty} [H_m(\chi)]^2 e^{\frac{-2\pi}{\Lambda}(\sqrt{\frac{\Lambda}{2\pi}}\chi)^2} \sqrt{\frac{\Lambda}{2\pi}} d\chi \\
&= \sqrt{\frac{\Lambda}{2\pi}} \int_{-\infty}^{\infty} H_m^2(\chi) e^{-\chi^2} d\chi \\
&= \sqrt{\frac{\Lambda}{2\pi}} 2^m m! \sqrt{\pi} \\
&= \sqrt{\frac{\Lambda}{2}} 2^m m! .
\end{aligned} \tag{VI.49}$$

Using this result in equation (VI.48), we continue

$$\begin{aligned}
P_{m,n} &= \frac{a_0^2}{1+\Omega^2} \left[\int_{-\infty}^{\infty} H_m^2 \left(\tilde{x} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{\frac{-2\pi\tilde{x}^2}{\Lambda}} d\tilde{x} \right] \left[\int_{-\infty}^{\infty} H_n^2 \left(\tilde{y} \sqrt{\frac{2\pi}{\Lambda}} \right) e^{\frac{-2\pi\tilde{y}^2}{\Lambda}} d\tilde{y} \right] \\
&= \frac{a_0^2}{1+\Omega^2} \left(\sqrt{\frac{\Lambda}{2}} 2^m m! \right) \left(\sqrt{\frac{\Lambda}{2}} 2^n n! \right) \\
&= \frac{a_0^2 \Lambda}{2(1+\Omega^2)} m! n! 2^m 2^n \\
&= P_{0,0} m! n! 2^m 2^n .
\end{aligned} \tag{VI.50}$$

Thus, we can normalize the power in each higher order mode to the power in the (0,0) mode simply dividing by $m! n! 2^m 2^n$. We can now define the mode normalized field as

$$A_{m,n} \equiv \frac{a_{m,n}}{\sqrt{m! n! 2^m 2^n}} . \tag{VI.51}$$

The next step is to determine the mode content of an unknown field. In our case, given an unknown field, $a(\tilde{\mathbf{r}}, \Omega)$, we know we can describe this field as a summation of numerous modes from our basis set where each mode has its own coefficient, $C_{m,n}$. We can write

$$a(\tilde{\mathbf{r}}, \Omega) = \sum_{m,n=0}^{\infty} C_{m,n} A_{m,n}(\tilde{\mathbf{r}}, \Omega) . \tag{VI.52}$$

To determine the mode content, we must determine the coefficients of interest. This is done through application of “Fourier’s trick” (a simple description can be

found in [66]). Start by multiplying both sides of equation (VI.52) by the complex conjugate of $A_{m',n'}$ and integrate over all space,

$$\begin{aligned}
\iint_{-\infty}^{\infty} a A_{m',n'}^* d\tilde{x} d\tilde{y} &= \iint_{-\infty}^{\infty} \sum_{m,n} C_{m,n} A_{m,n} A_{m',n'}^* d\tilde{x} d\tilde{y} \\
&= \iint_{-\infty}^{\infty} \sum_{m,n} C_{m,n} \frac{\sqrt{\frac{a_0^2}{1+\Omega^2}}}{\sqrt{m!n!2^m2^n}} H_m H_n e^{-2\pi\tilde{r}^2/\Lambda} e^{i\phi_{m,n}} \times \\
&\quad \frac{\sqrt{\frac{a_0^2}{1+\Omega^2}}}{\sqrt{m'!n'!2^{m'}2^{n'}}} H_{m'} H_{n'} e^{-2\pi\tilde{r}^2/\Lambda} e^{-i\phi_{m',n'}}. \tag{VI.53}
\end{aligned}$$

We pause to look at the imaginary exponential terms in equation (VI.53). Examining these terms,

$$e^{i\phi_{m,n}} e^{-i\phi_{m',n'}} = e^{i(\phi_{m,n} - \phi_{m',n'})},$$

we expand using equation (VI.42) and considering only the argument of the exponential,

$$\begin{aligned}
\phi_{m,n} - \phi_{m',n'} &= \frac{\pi\tilde{r}^2\Omega}{\Lambda} - (m+n+1)\tan^{-1}\Omega - \frac{\pi\tilde{r}^2\Omega}{\Lambda} + (m'+n'+1)\tan^{-1}\Omega \\
&= (m'-m+n'-n)\tan^{-1}\Omega, \\
e^{i\phi_{m,n}} e^{-i\phi_{m',n'}} &= e^{i[(m'-m+n'-n)\tan^{-1}\Omega]}. \tag{VI.54}
\end{aligned}$$

Since we know the HG polynomials are orthogonal, equation (VI.53) shows that $m' = m$ and $n' = n$ for a non-zero result then the imaginary exponent of equation (VI.54) is equal to unity. Using these observations, we continue solving for $C_{m,n}$, so that

$$\begin{aligned}
\iint_{-\infty}^{\infty} a A_{m',n'}^* d\tilde{x} d\tilde{y} &= \iint_{-\infty}^{\infty} \sum_{m,n} C_{m,n} \frac{\sqrt{\frac{a_0^2}{1+\Omega^2}}}{\sqrt{m!n!2^m2^n}} H_m H_n e^{-\frac{2\pi\tilde{r}^2}{\Lambda}} e^{i\phi_{m,n}} \times \\
&\quad \frac{\sqrt{\frac{a_0^2}{1+\Omega^2}}}{\sqrt{m'!n'!2^{m'}2^{n'}}} H_{m'} H_{n'} e^{-\frac{2\pi\tilde{r}^2}{\Lambda}} e^{-i\phi_{m',n'}} d\tilde{x} d\tilde{y} \\
&= C_{m,n} \left(\frac{a_0^2}{1+\Omega^2} \right) \left(\frac{1}{m!n!2^m2^n} \right) \times \\
&\quad \int_{-\infty}^{\infty} H_m^2 e^{-\frac{2\pi\tilde{x}^2}{\Lambda}} d\tilde{x} \int_{-\infty}^{\infty} H_n^2 e^{-\frac{2\pi\tilde{y}^2}{\Lambda}} d\tilde{y} \\
&= C_{m,n} \frac{a_0^2}{1+\Omega^2} \left(\frac{1}{m!n!2^m2^n} \right) \left(2^m m! \sqrt{\frac{\Lambda}{2}} \right) \left(2^n n! \sqrt{\frac{\Lambda}{2}} \right) \\
&= C_{m,n} \frac{a_0^2 \Lambda}{2(1+\Omega^2)}, \text{ so} \\
C_{m,n} &= \frac{2(1+\Omega^2)}{a_0^2 \Lambda} \iint_{-\infty}^{\infty} a A_{m,n}^* d\tilde{x} d\tilde{y}. \tag{VI.55}
\end{aligned}$$

To calculate the $C_{m,n}$, one need only calculate the overlap integral on the right-hand side of equation (VI.55) and multiply by a coefficient that is dependent on the scaling factors chosen. In general, $C_{m,n}$ is complex. If the intensity of an optical field is E^*E ($\propto a^*a$), which is a power per unit area and we have already normalized it so that the integral under the intensity curve is unity, then the percentage of energy or power in the (m, n) mode is described by $C_{m,n}^* C_{m,n}$.

C. CHARACTERIZING FEL FIELDS

With our basis set properly developed and the mathematical framework of the process established, we can now describe the methodology for analyzing an optical field and then apply it to real and proposed FEL fields as calculated using the NPS FEL simulations [1]. For our purposes, we ignore the source of the field and assume that an optical field has been generated and is provided for our analysis.

The first step is to normalize the optical intensity of the field. Since this is a simple scaling of the entire field, no modal decomposition information is lost, but it makes the computations and results much easier to understand and interpret. This

is accomplished in the usual manner, where a is the optical field provided,

$$1 = \iint_{-\infty}^{\infty} [ba^*(\tilde{x}, \tilde{y})] [ba(\tilde{x}, \tilde{y})] \, d\tilde{x} \, d\tilde{y}, \quad (\text{VI.56})$$

where solving for b gives the scaling required to ensure a normalized optical intensity.

The next step is to find the center of intensity. This is accomplished by calculating the first moment of the intensity (dropping the tildes for brevity),

$$\bar{x} = \frac{\int x |a(x, y)|^2 \, dx}{\int |a(x, y)|^2 \, dx}, \quad (\text{VI.57})$$

and similarly for \bar{y} .

The second-order moment of the field, which will be useful momentarily, is found by [63]

$$R^2 = 2 \frac{\iint |a(x, y)|^2 [(x - \bar{x})^2 + (y - \bar{y})^2] \, dx \, dy}{\iint |a(x, y)|^2 \, dx \, dy}. \quad (\text{VI.58})$$

We have yet to really describe the arbitrary scaling factors Ω and Λ introduced in equation (VI.41), except to say that we wish to choose them such that the number of modes required to describe a field is minimized. Although we cannot state initially what the values of these scaling factors are without some manipulation, we can bound their values based on their definitions.

We recall that Ω and Λ are defined

$$\begin{aligned} \Omega &\equiv \frac{\tau}{\tilde{z}_0}, \\ \Lambda &\equiv \pi \varrho^2. \end{aligned}$$

The curvature of the wavefront is tracked by Ω as it is related to the Gouy phase. Because the mode can be anywhere along its propagation path, there is no restriction on Ω 's value and it falls in the interval, $-\infty < \Omega < \infty$. A negative value for Ω indicates the field is propagating to a waist; a positive value indicates it is propagating away from a waist. On the other hand, we know that the mode waist radius must be some positive non-zero number by definition, so $\Lambda > 0$. As an upper bound on Λ , we

consider the second-order moment of the field of interest. This is the $1/e^2$ weighted radius of the intensity. We know that if the field of interest is composed only of the fundamental mode that the area of the mode using the $1/e$ point of the field is defined by Λ and this would be the same as πR^2 . For the same R^2 , if there are higher-order modes contained in the field of interest, it implies that the scaling factor must be something less than R^2 as the higher-order modes tend to push the intensity away from the centroid, as seen in Figure 117. Thus, we can establish πR^2 as the upper bound for Λ , putting Λ in the interval $0 < \Lambda \leq \pi R^2$.

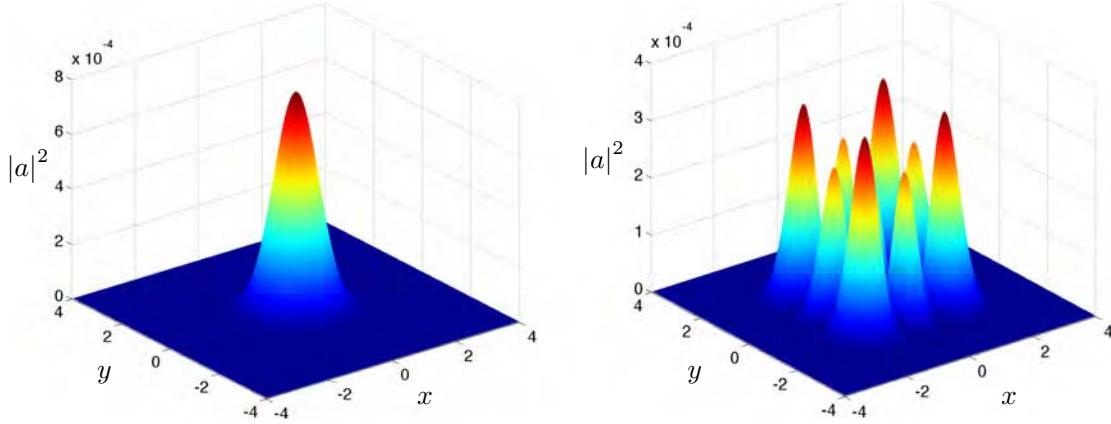


Figure 117. Intensity patterns for two modes generated with the following dimensionless parameters: $z_0 = 0.5$, $\tau = 0.625$ or $\Lambda = 4.0252$, $\Omega = 1.25$. Intensities are normalized so that the integrated intensity = 1. Transverse axis labels are dimensionless. Intensity is in arbitrary units. The left image is the intensity pattern of the fundamental mode with the $1/e^2$ radius, $\tilde{r} = 1.1319$. The right image is the intensity pattern of the 2-2 mode with the $1/e^2$ radius, $\tilde{r} = 2.5311$.

We must determine a discrimination method that can help us find the desired Ω and Λ in order to minimize the number of modes necessary to describe the field. To compare choices of the scaling factors, we establish a “coefficient area” to describe the number of modes and the amount of content in those modes. To do this, we find the mode with the greatest content and normalize the $|C_{m,n}|^2$ for this mode to 1. All the other $|C_{m,n}|^2$ are scaled appropriately and then all $|C_{m,n}|^2$ are summed.

To illustrate this concept, suppose we had a field composed of 50% (0,0), 25% (2,0) and 25% (0,2). This mode would have a coefficient area of 2.0. We can then perform additional decompositions with varying scaling factors and compare their coefficient areas and choose the least coefficient area as “better” as fewer contributing modes are necessary. Another simple illustration of this process is to consider a mode made up of only the fundamental. We can perform the decomposition for the correct scaling factors and again for incorrect scaling factors, and see how higher-order modes are populated contributing to increased coefficient area as demonstrated in Figure 118.

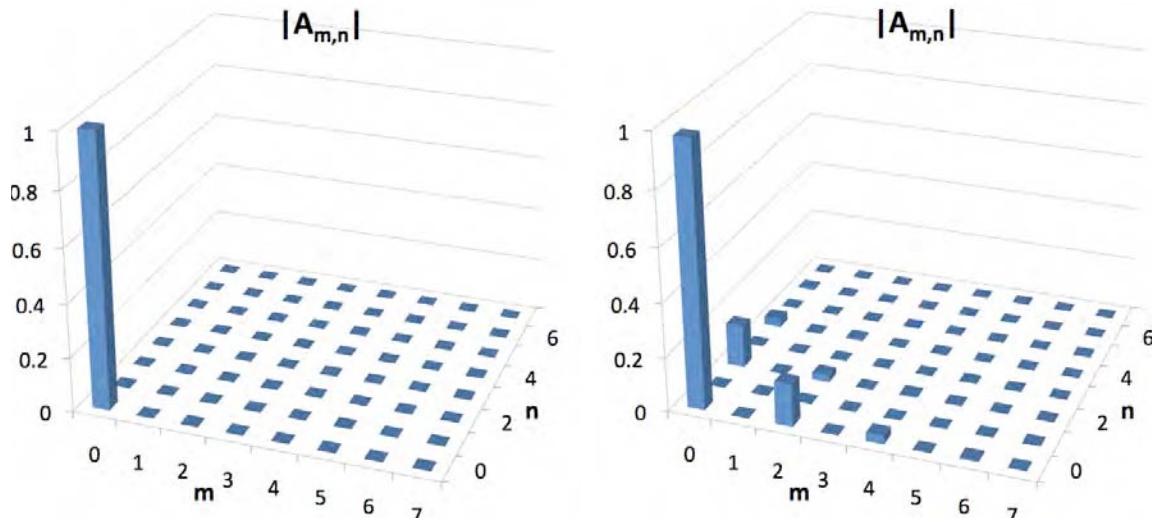


Figure 118. Modal composition histogram of a purely (0,0) mode (coefficient area = 1.0) against the correct basis, $\Omega = 1.25$ and $\Lambda = 4.0252$ (left) and a different scaling factor, $\Lambda = 3.0$, resulting in a coefficient area = 1.06 and finding 94.7% in the (0,0) mode and $\approx 2.5\%$ in the (0,2) and (2,0) modes (right) with the combination of the remaining modes constituting less than 0.3%. Histograms are plotted as $|C_{m,n}|$ to better show the higher-order modes.

We can use this coefficient area to build up a topology map showing the coefficient area found for many values of scaling factors. Figure 119 shows the topology when many scaling factors are examined for their resulting coefficient areas. Looking at this coefficient area map, we see that the behavior we expected: when the scaling factors are chosen carefully, there is a minimum value. Since we expect to be working

with fields that are primarily Gaussian as an assumption, we should expect to see coefficient area maps similar to Figure 119. The search area typically has a lower limit of about 10% of the πR^2 value estimated as the maximum Λ value discussed previously to minimize numerical effects that occur when the algorithm attempts to fit the field with a basis set with nearly zero cross-sectional area.

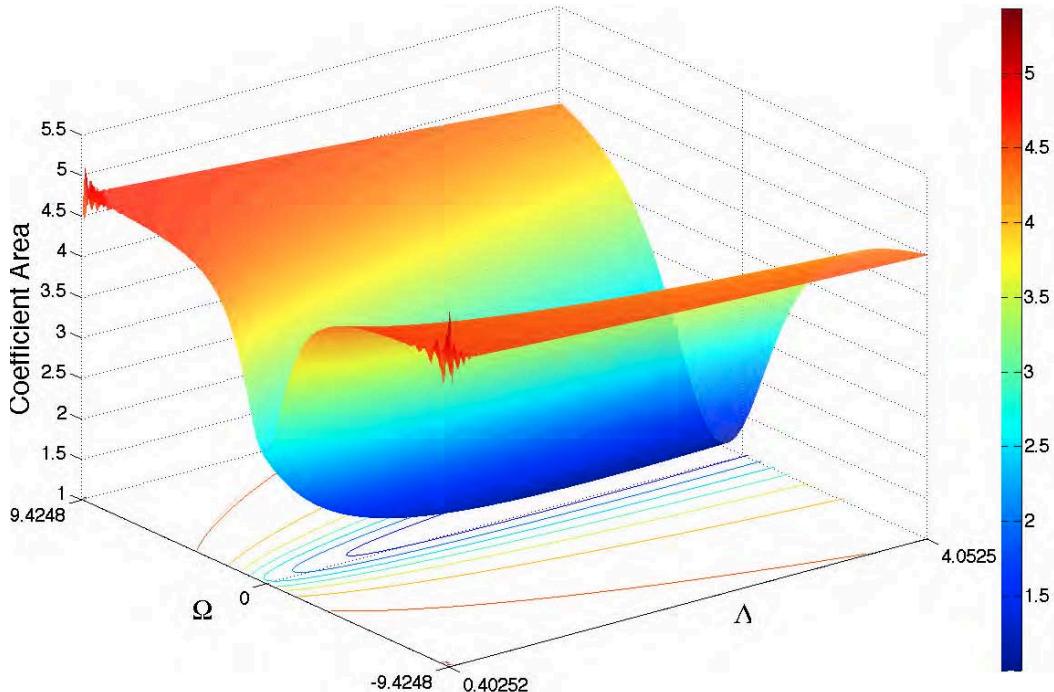


Figure 119. Topology of the coefficient area map for a purely $(0,0)$ mode with $\Lambda = 4.0252$ and $\Omega = 1.25$. The scaling factors Ω and Λ are varied over their respective ranges, demonstrating that the improper choice results in a greater coefficient area value.

Let us now see what happens when we generate coefficient area maps for modes that are nearly Gaussian to verify that there are no topological features that would give a simple minimization algorithm difficulties. We can see from Figure 120 that, in the nearly Gaussian case, there are no local minima that would cause difficulty for a simplistic minima-seeking algorithm. If the topology is complicated, as in the right hand subfigure of Figure 120, more robust or brute force search methodologies

must be used. In this more complicated case, the valley that rings the central peak would cause a minimum seeking algorithm to have to walk its way around once on the valley floor, greatly increasing computation time to achieve the minimum scaling factor solution.

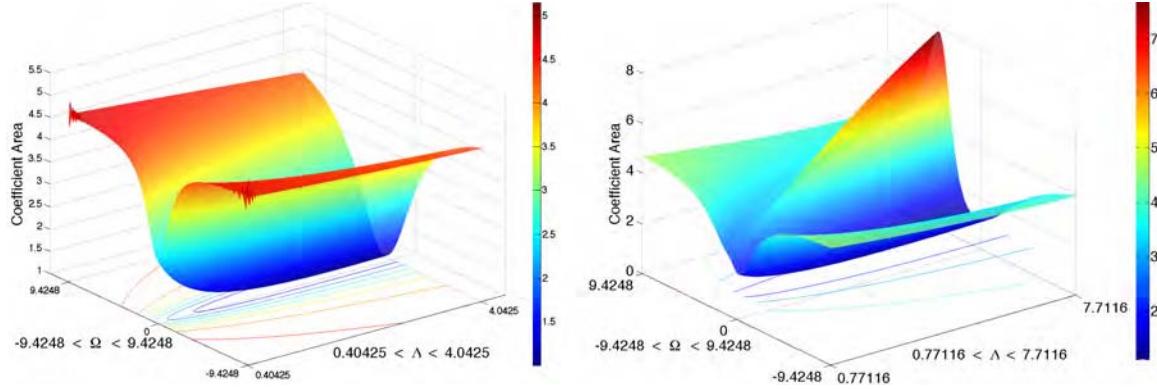


Figure 120. The coefficient area topology map on the left is generated from the modes shown in the right hand part of Figure 118 containing 95% (0,0) mode and 2.5% (2,0) and (0,2) modes. The map on the right is generated for a field of interest consisting of 50% (0,0), 20% (0,2), 20% (2,0), and 10% (2,2) modes.

Once we have minimized the number of modes required to represent a mode of interest, we must recognize that our algorithm, in actual practice, must truncate the number of modes used at some point. This truncation presents possible problems as we will discuss later when using this algorithm on actual FEL simulations. However, it is useful to have an independent method of measuring how well our coefficients represent the mode of interest. This can be done with a cross-correlation.

The cross-correlation is a tool used frequently in Fourier analysis and can be used as a figure of merit. The cross-correlation can be thought of as a matched-filter: it looks through an unknown signal (or field in this case) for a known signal (the output optical field from the FEL simulation). In mathematical terms, the one dimensional cross-correlation (\star), h , between two signals, f and g can be expressed

as

$$h(s) = f \star g = \int_{-\infty}^{\infty} f^*(x)g(s+x) \, dx , \quad (\text{VI.59})$$

where the asterisk indicates the complex conjugate and s is an offset variable allowing the signals to “shift” across one another. A graphical demonstration of the cross-correlation between two 3 by 3 arrays is shown in Figure 121 for illustrative purposes.

With large arrays, this direct process can become computationally expensive.

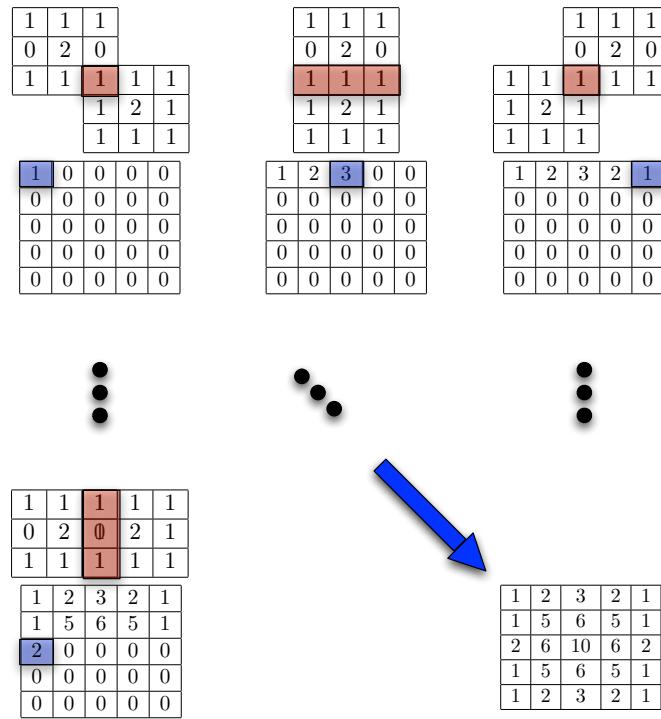


Figure 121. This is a simple demonstration of how a two-dimensional cross-correlation of two 3×3 arrays is computed in the spatial domain. Starting from the top left, the test image, T , is sequentially overlapped with the image of interest, a , as shown in brown. The sum of the element products are placed in the corresponding element of the resulting 5×5 cross-correlation array, shown in blue. This process continues until the last element of the resulting cross-correlation array is filled as shown in the bottom right-most array.

When large arrays are necessary, the Fourier convolution theorem offers us an alternative approach (see [65], for example). The convolution theorem states that the

Fourier transform, denoted as $\mathcal{F}\{ \cdot \}$, of a convolution in the spatial domain is the same as multiplication in the frequency domain. To illustrate this, let us examine a general convolution $(*)$ for which we desire to know h ,

$$\begin{aligned} h &= f * g \\ h(s) &= \int_{-\infty}^{\infty} f(x)g(s-x) \, dx \end{aligned} \quad (\text{VI.60})$$

now taking the Fourier transform of both sides,

$$\begin{aligned} \mathcal{F}\{h(s)\} = H(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)g(s-x)e^{-iks} \, ds \, dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)g(u)e^{-ik(u+x)} \, du \, dx \text{ where } u = s - x \\ &= \sqrt{2\pi} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u)e^{-iku} \, du \right) \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-ikx} \, dx \right) \\ &= \sqrt{2\pi} [\mathcal{F}\{g(u)\} \mathcal{F}\{f(x)\}] \\ H(k) &= \sqrt{2\pi} (G(k) F(k)) , \end{aligned} \quad (\text{VI.61})$$

we thus show that a convolution in the spatial domain is multiplication in the frequency domain. To get back to the spatial domain, we only need to take the inverse Fourier transform of $H(k)$. While this is illustrative, our process involves the correlation, not the convolution. The primary difference from equation (VI.60) is that $f(x) \rightarrow f^*(x)$ and $g(s-x) \rightarrow g(s+x)$. If we follow the same development as for the convolution, we prove the Wiener-Kinchin theorem and find that the cross-correlation in the spatial domain can be written

$$h(s) = \mathcal{F}^{-1}\{H(k)\} = \mathcal{F}^{-1}\left\{\sqrt{2\pi} (G(k) [F(k)]^*)\right\} , \quad (\text{VI.62})$$

where \mathcal{F}^{-1} is the inverse Fourier transform and $*$ denotes the complex conjugate. Since Fourier transforms have been highly optimized in computational codes, it is this approach that is used to determine the cross-correlation between the test field and the original field of interest.

Returning to our discussion of the cross-correlation in terms of the modal analysis algorithm, once the algorithm determines the correct scaling factors and the

coefficients, $C_{m,n}$, are determined, we can use them to generate a normalized test field, T . If the normalized test field, T , is an exact copy of a , the cross-correlation becomes an auto-correlation which has a maximum value of 1.0. If T is only a “near” match to a , then the cross-correlation will have some maximum value less than but “near” 1.0. It can be used to indicate whether the decomposition process worked well. This two-dimensional cross-correlation generates a surface that describes how well the test field and original field match up. The location of greatest cross-correlation value corresponds to the location of best match. Figure 122 shows the typical surface that results after one performs a two-dimensional cross-correlation.

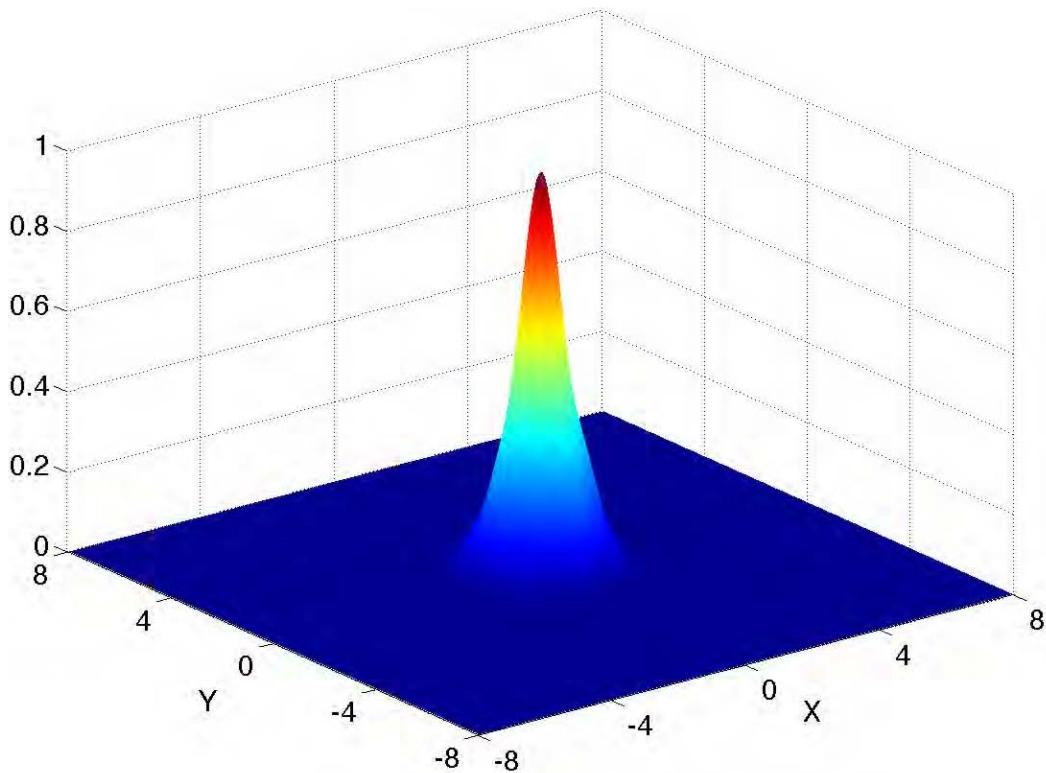


Figure 122. A two-dimensional auto-correlation surface of a pure Gaussian mode where the center of intensity is at the origin. If the mode were offset, the cross-correlation peak would correspond to the center of intensity. The expanded window (double the mode window of 8 in dimensionless units) is a result of the shifting-theorem used in the cross-correlation.

In a similar fashion to the coefficient area maps, we can generate maps of cross-correlation values where the scaling factor space can be mapped and the maximum two-dimensional cross-correlation value for each test field generated can be captured. Figure 123 shows cross-correlation maps generated for our two example fields used previously to generate the coefficient area maps in Figure 120. It might seem that it would be more efficient to simply use the cross-correlation as the discrimination factor and maximize this value. Computationally the cross-correlation is much more time consuming. For a 200 by 200 complex array (field of interest), the 200 by 200 point coefficient area maps in Figure 120 were completed in about 17 hours, or about 1.5 seconds per scaling factor considered. By contrast, the cross-correlation maps in Figure 123 took 22 hours total, or about 33% longer. Additionally, if we compare Figures 120 and 123, the topology for finding an extremum appears to have more discernible features in the coefficient area case. Thus, our decision to use the coefficient area as the primary discriminator for choosing appropriate scaling factors is validated on two accounts. The coefficient area approach has less of a computational cost than the cross-correlation and the topology found with the coefficient area map (of which, we have only shown two cases, but was generally noted from a much larger sample size) has more topological features allowing for better extremum discrimination.

The approach is now fully described and we can turn our attention to fields generated using real FEL parameters. For these studies, the NPS 4D FEL amplifier simulation, described in [67], was used to simulate four FELs based on the best parameters available, shown in Table 9. These particular FELs were chosen, primarily because the published data allowed for generation of the necessary input parameters for simulation. The Brookhaven National Laboratory (BNL) designs include a proposed untapered undulator design and moderate decreasing linear taper used in their Source Development Lab (SDL) and the Los Alamos National Laboratory (LANL) designs provide a gentler decreasing taper and a step taper for analysis.

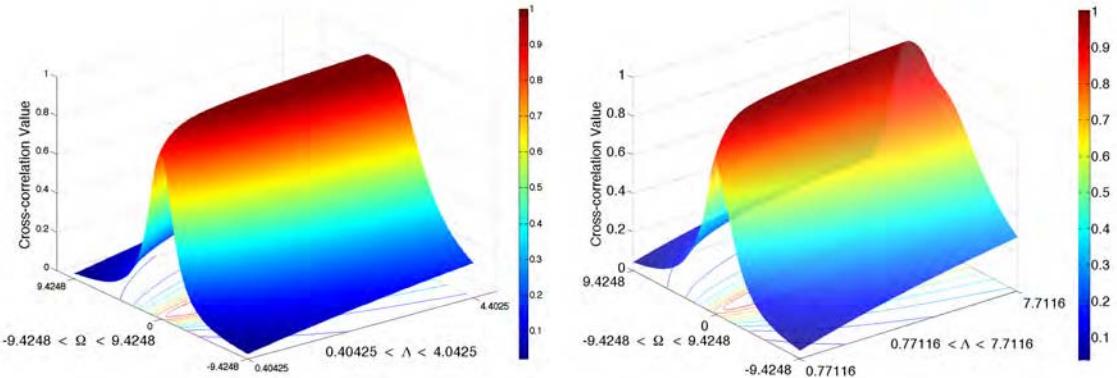


Figure 123. The two cross-correlation maps are those for the same modes as shown in the coefficient area maps in Figure 120. The map on the left is for a nearly Gaussian field: 95 % (0,0) and 2.5% (2,0) and (0,2) modes. The map on the right is for a field containing 50% (0,0), 20% (0,2), 20% (2,0), and 10% (2,2) modes.

The general algorithm described above was coded to perform the necessary calculations in a program separate from the NPS 4D FEL simulation code. The NPS FEL code provided the output of the optical field, real and imaginary parts, and necessary parameters for scaling for use as input into the analysis code. All optical analysis was performed just after the interaction region (at the end of the undulator) using up to the 7th order HG mode in each dimension (a total of 64 modes). The modal analysis results are tabulated in Table 10.

For the four cases presented in Table 10, we find the amplifier FELs generate optical fields primarily in the fundamental mode. Using the minimum number of modes to describe the optical field, we see that the FELs are primarily composed of the fundamental mode, with cross-correlation values near 0.99, indicating excellent agreement between the simulation's output data and a comparison field generated using the modes found from this analysis approach. The BNL SDL FEL is an interesting case as it was the only FEL that the analysis tool had problems decomposing. Regardless of the search starting point, the algorithm would find a minimum that

Table 9. FEL parameters used for HG decomposition.

	BNL SDL [68]	BNL Proposed [69]	LANL Proposed [70] step taper	LANL Proposed [70] linear taper
N	256	120	110	110
λ_0	3.9 cm	3.275 cm	2.18 cm	2.18 cm
K_{rms}	0.78	0.7	1.2	1.2
E	102 MeV	80 MeV	81 MeV	81 MeV
q	0.35 nC	1.4 nC	1 nC	1 nC
t_b	1 ps	2.8 ps	1 ps	1 ps
$\frac{\Delta B}{B}$	-5%	0%	-2.5%	-8.0%

had a cross-correlation value of around 0.73, indicating a poor match to the original field. It became apparent that the field of interest required additional higher-order modes to properly decompose the field. By doubling the number of modes in use in both dimensions (quadrupling the total number of modes), the algorithm was able to successfully find a set of scaling factors that generated a high cross-correlation value, which is reported in Table 10.

If we examine the FEL simulation for this particular case, it is evident that there are possible reasons for needing more modes. Figures 124 and 125 show the relevant output from the FEL simulation. From Figure 124, we can see that there is significant structure surrounding the central lobe. This structure becomes significantly more prominent as the mode is allowed to propagate further, although the mode content remains the same. In Figure 125, we see that the FEL reached saturation before even half the undulator was traversed as the gain stabilized at about $\tau = 0.5$. Saturation effects can drive power into higher-order modes in FELs at the expense of the fundamental mode. Thus, we were justified in extending the coefficient series to higher-order modes, because if there had been no higher order mode content, there would have been no improvement in the the cross-correlation value for the minimum coefficient area solution. Had we chosen a basis set that assumed

Table 10. Modal analysis results of four amplifier FELs. The five greatest contributing modes ($\geq 1\%$) and the percentage of their composition are shown. Cross-correlation values are reported as a figure of merit (FOM) for a field generated from the decomposed mode coefficients and corresponding scaling factors Ω and Λ , with the simulation output field. (* Used $14 \times 14 = 196$ modes to perform decomposition.)

	Modes	% Composition	Ω	Λ	FOM
BNL SDL*	(0,0)	32.3%	1.053	2.374	0.983
	(2,0)	15.0%			
	(0,2)	15.0%			
	(0,4)	6.4%			
	(4,0)	6.3%			
BNL proposed	(0,0)	95.3%	1.129	1.138	0.996
	(2,2)	1.0%			
	(0,4)	1.3%			
	(4,0)	1.3%			
LANL Step	(0,0)	99.4%	3.675	5.763	0.991
LANL Linear	(0,0)	94.4%	2.545	3.748	0.987
	(0,2)	1.5%			
	(2,0)	1.4%			
	(0,4)	1%			
	(4,0)	1%			

axisymmetric modes, such as the Laguerre-Gaussian polynomials, fewer modes may have been required to decompose this particular field of interest, but the ability to handle non-axisymmetric beams would have been lost.

This methodology provides researchers and designers another method of measuring the quality of their FEL designs. By monitoring the mode content of the output optical field, the propagation characteristics can be predicted and optimized for the desired application. We have developed a robust and adaptable algorithm for decomposing arbitrary FEL optical fields as well as a well-known figure of merit for determining the success of the decomposition. Additionally, the decomposition method has provided confirmation that many amplifier FELs operate nearly in the fundamental mode regardless of their undulator design and that the mode content observed reflects well known FEL interaction phenomena such as saturation.

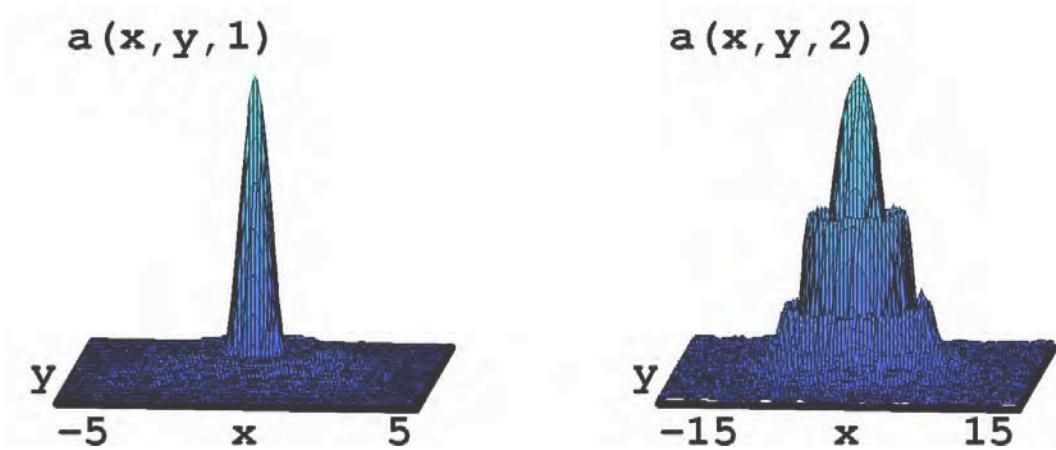


Figure 124. The dimensionless optical field amplitude, $|a(x, y)|$. The image on the left is the field at the end of the undulator. The field on the right is the field after propagating another undulator length. Intensities are in arbitrary units and do not represent the same scale between pictures.

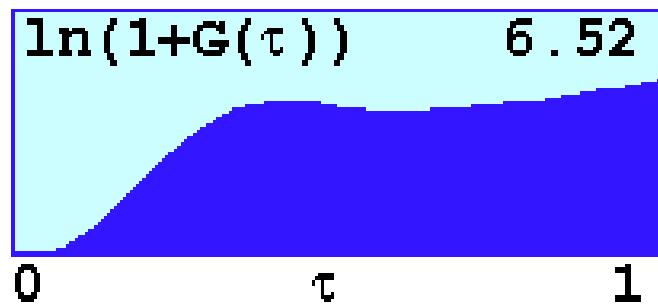


Figure 125. The gain, G , experienced by the optical field as it traverses the undulator of the BNL SSDL FEL.

VII. CONCLUSION

In this work, we examined a free electron laser (FEL) from start to end. We started with an examination of some of the cathode types available for use in an electron source. Electron sources in the form of guns, boosters, and injectors were then discussed with simulation results presented for many different configurations and electron beam shapes. From simulation, we moved to discussion about the design and construction of the first superconducting radio frequency (SRF) gun in the United States. Experimental results showing the performance of the first prototype SRF gun were presented and compared with simulations showing that the gun is currently capable of generating 10% of the desired bunch charge at 25% design energy in only its first beam generation experiment.

In the DC gun, the thermionic cathode has proven why it is the workhorse of electron beam generation in many applications. Unfortunately, the basic thermionic cathode's applicability to short pulse regimes like those found in SRF guns is limited without additional beam line space being used for buncher cavities and compressor chicanes. If the desire is to decrease the overall FEL footprint, some other mechanism for pulsing the thermionic cathode, such as photogating, will be necessary, as has been shown to be practicable in our experiments.

Bare metal cathodes have been proven in many other systems and demonstrated in the NPS SRF gun. They are, in principle, robust and do not require a great deal of “care and feeding” to maintain performance. There are preparation schemes for metal cathodes that promise an increase in quantum efficiency (QE) that could drive the gun bunch charge to levels approaching 1 nC per bunch. Alkali coated semiconductor cathodes currently in use in DC guns could also migrate into SRF guns. Unfortunately, not much is known about the effects of high QE materials, such as cesium, migrating onto the interior surfaces of an SRF cavity. SRF cavity cleanliness requirements are already extremely stringent, but as improved models of

the NPS SRF gun are built, previous editions could be used as test platforms for the cesiated cathodes without worry of impact to ongoing NPS experimental programs. Once the gun fields in the cathode region approach those necessary for turn-on of field emission arrays, the NPS SRF gun is an ideal platform for testing them in a realistic gun environment, especially with the ability to adjust the cathode position and the impinging electric fields almost at will.

The DC gun received from Stanford was originally envisioned as the electron source for the NPS FEL and as a test bed for cathode studies. As a test bed, it has proven to be capable, but the full use of the device has been hindered by local radiation policy issues forcing a change in experimental direction that has impacted data collection from this useful tool. As an exercise in returning a significant piece of a decommissioned linear accelerator to service, the lessons learned in returning the DC gun to service and processing it back to near peak operating voltage were invaluable to the experimental group at NPS. It also serves as a case study for how rugged DC guns and thermionic cathodes are, even when stored in less than ideal conditions.

Simulation studies of the NPS SRF cavity in the booster configuration showed that the booster design is successful. When coupled with the former Stanford DC gun, simulations indicate that transport of the beam through the booster cavity and into a merge can be accomplished, but would require beam focusing prior to entry into the booster, otherwise charge transport would suffer. Coupling the booster to the DC gun as an injector would be less than ideal due to the (i) extended beam line necessary to provide focusing to counter space charge effects while the beam is still at low energy, as well as (ii) the addition of a buncher cavity to compress the bunch sufficiently to correct for longitudinal beam growth.

With the addition of a cathode and stalk to place the cathode, the NPS SRF booster cavity can be transformed into an electron source in its own right. The prototype cavity has demonstrated through simulation and experiment that it would

be a suitable gun. Transport acceptance is wide, allowing for nearly 180° of relative phase for which the beam can be fully transported out of the gun. Normalized emittances of 7 mm-mrad from 1 nC bunch charges at the design energy of 1.2 MeV have been simulated but await verification by experiment. The coaxial coupler has proven less problematic as experience with it has increased, but it still requires a significant drift space for insertion. The feared aperture impact of the coupler on the electron beam proved to be less of an issue than expected; however, it may be advantageous to consider other methods of coupling RF to the cavity that would allow for a more compact device. Experimentally, we verified the deleterious impact of the cathode stalk on the cavity fields, but we have not been able to quantify it as yet. It does indicate that a different cathode insertion system may be warranted.

As part of the merger studies, a double quarter wave (or highly re-entrant pillbox) cavity is proposed as a booster. When coupled with the NPS SRF gun, it could constitute a > 3 MeV injector for the NPS FEL linear accelerator. The double quarter wave booster, when optimized for small spot size or best emittance, was capable of delivering a suitable beam for injection into a linear accelerator in simulation. This booster design has the advantage of requiring only minimal development as the drawings and procedures for production would be based on a proven cavity design. To overcome problems with cooling of the superconducting solenoid in the NPS SRF cavity, the solenoids should be immersed in the liquid helium cryostat encompassing both the gun and booster. Both the long and short versions of the booster showed little emittance growth and were successful in transporting 1 nC bunches without issues.

Shifting from experimental to theoretical studies, an analysis tool for FEL optical fields has been developed. The methodology is based upon Hermite-Gaussian decomposition, which has proven useful in the analysis of oscillator FELs. Unlike the oscillator, in which the basis set scaling factors can be determined from the oscillator cavity, amplifier FELs have no unique factors that would fix the scaling for

decomposition. Under the assumption that the simplest answer is the best answer, a methodology for choosing scaling factors such that the number of Hermite-Gaussian polynomial terms necessary to decompose an optical field of interest was developed and implemented in the NPS FEL4D simulation code. Using the simulation code and parameters from the literature, the FEL optical fields of operating and proposed amplifiers were analyzed and demonstrated to operate primarily in the fundamental mode. By using a decomposition methodology similar to those used by other laser types, it allows for a more judicious side-by-side comparison between FELs and traditional lasers when discussing propagation characteristics.

In the future, as the NPS FEL4D code is tied to particle accelerator simulation codes and the NPS FEL is built, opportunities to examine an entire FEL system in simulation and experiment may become possible. It may soon be possible for a researcher to propose a system component, model it in a particle accelerator code such as GPT, perform FEL simulations, and predict the optical transport of the laser light and then perform the experiment for validation. Exciting times are in store for future students and researchers as the NPS FEL facilities and capabilities continue to grow.

LIST OF REFERENCES

- [1] S. P. Niles, J. Blau, and W. B. Colson, “Hermite-Gaussian decomposition of free electron laser optical fields,” *Physical Review Special Topics - Accelerators and Beams*, vol. 13, p. 030702, August 2010.
- [2] H. J. Reich, *Theory and Applications of Electron Tubes*, 2nd ed. New York: McGraw-Hill Book Company, 1944.
- [3] Bureau of Naval Personnel, Ed., *Electricity*, NavPers 10622-B ed. United States Government Printing Office, 1951.
- [4] M. Riser, *Theory and Design of Charged Particle Beams*, 2nd ed. Weinheim: Wiley-VCH, 2008.
- [5] R. Forman and D. H. Smith, “Thermionic cathode life-test studies,” *IEEE Transactions on Electron Devices*, vol. 26, no. 10, pp. 1567–1572, October 1979.
- [6] D. R. Lide, Ed., *Handbook of Chemistry and Physics*, 84th ed. New York: CRC Press, 2003.
- [7] W. P. Wang et al., “Diamond vacuum field emission devices,” *Diamond and Related Materials*, vol. 13, pp. 1944–1948, 2004.
- [8] J. W. Lewellen and J. Noonan, “Field-emission cathode gating for RF electron guns,” *Physical Review Special Topics – Accelerators and Beams*, 2005.
- [9] C. Hernandez-Garcia et al., “A high average current DC GaAs photocathode gun for ERLs and FELs,” *Proceedings of the 2005 Particle Accelerator Conference*, 2005.
- [10] I. V. Bazarov and C. K. Sinclair, “High brightness, high current injector design for the Cornell ERL prototype,” *Proceedings of the 2003 Particle Accelerator Conference*, 2003.
- [11] A. M. M. Todd et al., “State-of-the-art electron guns and injector designs for energy recovery linacs (ERL),” *Proceedings of the 2005 Particle Accelerator Conference*, 2005.
- [12] H. Padamsee, J. Knobloch, and T. Hays, *RF Superconductivity for Accelerators*. Hoboken: John Wiley and Sons, 1998.
- [13] D. Janssen et al., “First operation of a superconducting RF-gun,” *Nuclear Instruments in Physics Research A*, pp. 314–317, 2003.

- [14] C. Bohnet. Erste supraleitende kanone für beschleuniger im einsatz. [first superconducting accelerator cannon in use]. [Online]. Accessed March 28, 2010: <http://www.fzd.de/db/Cms?pOid=25818&pNid=473>
- [15] A. Arnold et al., “1st RF-measurements @ 3.5-cell SRF-photo-gun cavity in Rossendorf,” *Proceedings of FEL 2006*, pp. 567–570, 2006.
- [16] J. Teichert et al., “Progress of the Rossendorf SRF gun project,” *Proceedings of the 27th International Free Electron Laser Conference*, 2005.
- [17] J. W. Lewellen, W. B. Colson, S. P. Niles, and T. Smith, “Status of the NPS free-electron laser,” *Proceedings of LINAC08*, 2008.
- [18] J. W. Lewellen et al., “The NPS-FEL injector upgrade,” *Proceedings of LINAC08*, 2008.
- [19] D. J. Griffiths, *Introduction to Electrodynamics*, 3rd ed. Upper Saddle River: Prentice-Hall, 1999.
- [20] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. Hoboken: John Wiley and Sons, 1999.
- [21] C. M. Lyneis, H. A. Schwettman, and J. P. Turneaure, “Elimination of electron multipacting in superconducting structures for electron accelerators,” *Applied Physics Letters*, vol. 31, no. 8, pp. 541–543, October 1977.
- [22] H. Padamsee and A. Joshi, “Secondary electron emission measurements on materials used for superconducting microwave cavities,” *Journal of Applied Physics*, vol. 50, pp. 1112–1115, 1979.
- [23] W. P. Kang, A. Wisitsora-at, J. L. Davidson, and M. Howell, “Micropattern-gated diamond field emitter array,” *Journal of Vacuum Science and Technology B*, Mar/Apr 1998.
- [24] D. H. Dowell, R. F. Boyce, C. Carr, G. J. Collet, and R. E. Kirby, “Conceptual design of the LCLS 120 Hz gun load lock,” Stanford Linear Accelerator Center, Tech. Rep. SLAC-TN-05-018, April 2004.
- [25] Q. Zhao, T. Srinivasan-Rao, and M. Cole, “Tests of niobium cathode for the superconducting radio frequency gun,” *Proceedings of the 2003 Particle Accelerator Conference*, pp. 2047–2049, 2003.
- [26] T. Rao et al., “Photoemission studies on BNL/AES/JLAB all niobium, superconducting RF injector,” *Proceedings of the 2005 Particle Accelerator Conference*, pp. 2556–2558, 2005.

- [27] T. Srinivasan-Rao, J. Fischer, and T. Tsang, “Photoemission studies on metals using picosecond ultraviolet laser pulses,” *Journal of Applied Physics*, vol. 69, pp. 3291–3296, March 1991.
- [28] Ph. Bernard et al., “Superconducting niobium sputter-coated copper cavities at 1500 MHz,” *Proceedings of EPAC 1992*, pp. 1269–1271, 1992.
- [29] The MuShield Company. About M μ Metal. [Online]. Accessed May 5, 2010: www.mumetal.com/about_mumetal.html
- [30] K. Halbach and R. F. Holsinger, “SUPERFISH - a computer program for evaluation of RF cavities with cylindrical symmetry,” *Particle Accelerators 7*, pp. 213–222, 1976.
- [31] J. H. Billen and L. M. Young, “Poisson SUPERFISH,” Los Alamos National Labs, Tech. Rep. LA-UR96-1834, 2006.
- [32] J. H. Billen, “Parmela,” Los Alamos National Labs, Tech. Rep. LA-UR-96-1835, December 2005.
- [33] S. B. van der Geer and M. J. de Loos, *General Particle Tracer User Manual*, Pulsar Physics.
- [34] S. B. van der Geer, O. J. Luiten, M. J. de Loos, G. Pöplau, and U. van Rienan, “3D space-charge model for GPT simulations of high-brightness electron bunches,” *Institute of Physics Conference Series*, no. 175, p. 101, 2005.
- [35] O. J. Luiten, S. B. van der Geer, M. J. de Loos, F. B. Kiewiet, and M. J. van der Wiel, “How to realize uniform three-dimensional ellipsoidal electron bunches,” *Physical Review Letters*, vol. 93, no. 9, August 2004.
- [36] K. L. Jensen, D. W. Feldman, M. Virgo, and P. G. O’Shea, “Measurement and analysis of thermal photoemission from a dispenser cathode,” *Physical Review Special Topics – Accelerators and Beams*, vol. 6, no. 083501, 2003.
- [37] K. L. Jensen, D. W. Feldman, and P. G. O’Shea, “The quantum efficiency of dispenser photocathodes: Comparison of theory to experiment,” *Applied Physics Letters*, vol. 85, no. 22, pp. 5448–5450, November 2004.
- [38] Y. E. Sun, J. W. Lewellen, and D. W. Feldman, “Photothermal cathode measurements at the advanced photon source,” *Proceedings of LINAC 2006*, pp. 349–351, 2006.
- [39] Communications & Power Industries. Electron sources. [Online]. Accessed June 29, 2010: <http://www.cpii.com/docs/datasheets/131/Electron-Sources-Eimac-Data%20Sheet-Eimac.pdf>

- [40] CPI, “Care and feeding of power grid tubes,” Communications and Power Industries, Tech. Rep., 2003.
- [41] J. Jimenez, “A systematic cathode study—activation of thermionic cathode and measuring cesium evaporation from a dispenser photocathode,” Master’s thesis, Naval Postgraduate School, 2010.
- [42] S. P. Niles et al., “NPS prototype superconducting 500 MHz quarter-wave gun update,” *Proceedings of the 32nd International Free Electron Laser Conference*, preprint, 2010.
- [43] J. C. Slater, *Microwave Electronics*. New York: Van Nostrand, 1950.
- [44] P. Goudket et al., “Comparison of stretched-wire, bead-pull and numerical impedance calculations on 3.9 GHz dipole cavities,” *Proceedings of EPAC 2008*, 2008.
- [45] D. C. Carey, K. L. Brown, and F. Rothacker, “Third-order TRANSPORT with MAD input a computer program for designing charged particle beam transport systems,” Stanford Linear Accelerator Center, Tech. Rep. SLAC-R-530, October 1998.
- [46] S. Y. Lee, *Accelerator Physics*, 2nd ed. Hackensack: World Scientific, 2004.
- [47] J. Knobloch, H. Muller, and H. Padamsee, “Design of a high-speed, high-resolution thermometry system for 1.5-GHz superconducting radio-frequency cavities,” *Review of Scientific Instruments*, vol. 65, no. 11, pp. 3521–3527, November 1994.
- [48] R. Geng, “Protocol for performing gas helium processing of SRF cavities in SRF group at Cornell,” Cornell University, Tech. Rep. LEPP-SRFP-RG01, December 2003.
- [49] O. Brunner, A. Butterworth, G. Cavallari, N. Hilleret, J. Jimenez, and J. Tuckmantel, “First experience with in situ helium processing of the LEP superconducting modules,” *Proceedings of the 1997 Workshop on RF Superconductivity*, pp. 133–137, 1997.
- [50] W. B. Colson, “Free electron laser theory,” Ph.D. dissertation, Stanford University, 1977.
- [51] S. P. Niles, “High energy laser applications in a surface combatant: Terminal phase theater ballistic missile defense, low atmosphere propagation, and free electron laser gain,” Master’s thesis, Naval Postgraduate School, 2005.
- [52] C. A. Brau, *Free-Electron Lasers*. Boston: Academic Press, 1990.

- [53] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, no. 4736, pp. 493–494, August 1960.
- [54] J. M. J. Madey, "Stimulated emission of bremsstrahlung in a periodic magnetic field," *Journal of Applied Physics*, vol. 42, no. 5, pp. 1906–1913, April 1971.
- [55] L. R. Elias, W. M. Fairbank, J. M. J. Madey, H. A. Schwettman, and T. I. Smith, "Observation of stimulated emission of radiation by relativistic electrons in a spatially periodic transverse magnetic field," *Physical Review Letters*, vol. 36, no. 13, pp. 717–720, March 1976.
- [56] W. B. Colson, J. Blau, K. Cohn, J. Jimenez, and R. Pifer, "Free electron lasers in 2009," *Proceedings of FEL 2009*, pp. 591–595, 2009.
- [57] S. Benson et al., "First lasing of the Jefferson Lab IR demo FEL," *Nuclear Instruments in Physics Research A*, no. 429, pp. 27–32, 1999.
- [58] C. Behre et al., "First lasing of the IR upgrade FEL at Jefferson Lab," *Nuclear Instruments in Physics Research A*, no. 528, pp. 19–22, 2004.
- [59] Office of Naval Research. Free electron laser program. [Online]. Accessed July 11, 2010: <http://www.onr.navy.mil/Science-Technology/Departments/Code-35/All-Programs/air-warfare-352/Free-Electron-Laser.aspx>
- [60] Thomas Jefferson National Accelerator Facility. Free-electron laser description. [Online]. Accessed July 11, 2010: <http://www.jlab.org/FEL/felspecs.html>
- [61] Centre Laser Infrarouge d'Orsay. What is a free electron laser? [Online]. Accessed July 11, 2010: http://clio.lcp.u-psud.fr/clio_eng/FELrad.html
- [62] A. E. Siegman, *Lasers*. Sausalito: University Science Books, 1986.
- [63] T. S. Ross and W. P. Latham, "Appropriate measures and consistent standard for high energy laser beam quality," *Journal of Directed Energy*, vol. Summer, pp. 22–58, 2006.
- [64] R. Vigil, "Hermite-Gaussian modes and mirror distortions in the free electron laser," Master's thesis, Naval Postgraduate School, June 2006.
- [65] K. F. Riley, M. P. Hobson, and S. J. Bence, *Mathematical Methods for Physics and Engineering*, 2nd ed. Cambridge: Cambridge University Press, 2002.
- [66] D. J. Griffiths, *Introduction to Quantum Mechanics*, 2nd ed. Upper Saddle River: Prentice-Hall, 1995.
- [67] J. Blau, D. Burggraaff, T. Y. Voughs, and W. B. Colson, "Simulations of high power-FEL amplifiers," *Proceedings of the 28th International Free Electron Laser Conference*, 2006.

- [68] T. Watanabe et al., “An experimental test of superradiance in a single pass seeded FEL,” *Proceedings of the 27th International Free Electron Laser Conference*, 2005.
- [69] I. Ben-Zvi, D. Kayran, and V. Litvinenko, “High average power optical FEL amplifiers,” *Proceedings of the 27th International Free Electron Laser Conference*, 2005.
- [70] D. C. Nguyen, H. P. Freund, and W. B. Colson, “A scalloped electron beam free-electron laser,” *Proceedings of the 28th International Free Electron Laser Conference*, 2006.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, VA
2. Dudley Knox Library
Naval Postgraduate School
Monterey, CA
3. Dean of Research
Naval Postgraduate School
Monterey, CA
4. Chairman, Physics Department
Naval Postgraduate School
Monterey, CA
5. Quentin Saulter
Office of Naval Research
Arlington, VA
6. Dr. Terry Grimm
Niowave, Inc.
Lansing, MI
7. Professor William B. Colson
Naval Postgraduate School
Monterey, CA
8. Professor John Lewellen
Naval Postgraduate School
Monterey, CA
9. Professor John Harris
Naval Postgraduate School
Monterey, CA